How large are the non-specific effects of acupuncture? A meta-analysis of randomized controlled trials

Klaus Linde[1]*, MD; Karin Niemann[1], cand.med.; Antonius Schneider[1], MD; Karin Meissner[1,2], MD

[1]Institute of General Practice, Technische Universität München, Wolfgangstr.8, 81667 Munich, Germany

[2]Institute of Medical Psychology, Ludwig-Maximilians-University, 80336 Munich, Germany

*Corresponding author: Klaus Linde

Email addresses:

KL: klaus.linde@lrz.tum.de

KN: karin.niemann@gmx.net

AS: antonius.schneider@lrz.tum.de

KM: karin.meissner@med.lmu.de

Abstract

Background

While several recent large randomized trials found clinically relevant effects of acupuncture over no treatment or routine care, blinded trials comparing acupuncture to sham interventions often reported only minor or no differences. This raises the question whether (sham) acupuncture is associated with particularly potent non-specific effects. We aimed to investigate the size of non-specific effects associated with acupuncture interventions.

Methods

Medline, Embase, Cochrane Central Register of Controlled Clinical Trials, and reference lists were searched up to April 2010 to identify randomized trials of acupuncture for any condition including both a sham and a no acupuncture control group. Data were extracted by one reviewer and verified by a second. Pooled standardized mean differences were calculated using a random effects model with the inverse variance method.

Results

37 trials with a total of 5754 patients met the inclusion criteria. The included studies varied strongly regarding patients, interventions, outcome measures, methodological quality and effect sizes reported. Among the 32 trials reporting a continuous outcome measure the random effects standardized mean difference between sham acupuncture and no acupuncture groups was -0.45 (95%CI -0.57, -0.34; I^2 = 54%, Egger's test for funnel plot asymmetry p = 0.25). Trials with larger effects of sham over no acupuncture reported smaller effects of acupuncture over sham intervention than trials with smaller non-specific effects (β = -0.39, p = 0.029).

Conclusions

Sham acupuncture interventions are often associated with moderately large non-specific effects, which could make it difficult to detect small additional specific effects. Compared to

inert placebo interventions effects associated with sham acupuncture might be larger, which would have considerable implications for the design and interpretation of clinical trials.

Background

In recent years there is increasing evidence from large randomized trials and systematic reviews showing that patients receiving acupuncture report better outcomes than patients receiving no treatment or usual care only (for example [1-2]). A large trial on low back pain [1] and a meta-analysis of migraine trials [2] even found superiority over guideline-oriented conventional care. At the same time many recent high quality trials comparing true acupuncture with a sham acupuncture intervention found only minor or even no differences (see [2-5] for systematic reviews). The interpretation of this evidence is controversial. Some authors argue that the effects over no treatment and usual care are only due to usual placebo effects and bias [3]. Some authors argue that most sham acupuncture interventions are physiologically active [4-5] and others that sham acupuncture interventions might be associated with particularly potent non-specific or placebo effects [6-7].

Treatment effects are considered specific if they are attributable solely to the – according to the theory of the mechanism of action - characteristic component of an intervention [8-9]. Effects which are associated with the incidental elements of an intervention are considered non-specific effects (syn. placebo effects). Non-specific effects are mostly thought to be due to psychobiological processes triggered by the overall therapeutic context [10]. They have to be distinguished from the natural course of disease, regression to the mean, the effects of being in a study, co-interventions and, as far as possible, from reporting and other biases [11-12]. The total effect of an intervention consists of both specific and non-specific effects [13].

Separating characteristic and incidental elements of an intervention is straightforward in pharmacology but is difficult in other interventions such as psychotherapy [14]. Acupuncture involves the insertion and manipulation of needles into defined points of the body. While a variety of mechanistic models exist the exact mechanism of action is unclear [15]. This makes it difficult to devise a placebo intervention which is both inert and indistinguishable, and reliably separates specific and non-specific effects. The frequent use of the word sham

intervention instead of placebo partly reflects this problem. Sham interventions in clinical trials of acupuncture typically vary from “true” acupuncture in one or both of the following aspects [16]: location of points (for example, stimulation of non-indicated points or outside of known points) and skin penetration (for example, use of fixed telescope “placebo” needles with a blunt tip). If some or most of these sham interventions should indeed be physiologically active such trials would not compare acupuncture to a placebo but to an active intervention making it more difficult to detect significant differences.

This problem would also apply if (sham) acupuncture would be associated with more potent placebo effects than other interventions. Both invasive and non-invasive sham acupuncture interventions exert (like true acupuncture) mild painful stimuli. It has been hypothesized that such interventions might trigger enhanced placebo effects by simultaneously acting on a sensory, cognitive and emotional level [7]. There is also evidence that the same sham acupuncture intervention can have quite different effects when provided in different contexts [17]. Placebo research indicates that in many situations the therapeutic context associated with an intervention matters more than the placebo intervention itself [20-21]. The therapeutic context is depending not only on the specific therapeutic ritual applied but also on experiences, attitudes and preferences of patients and providers, on the patient-provider interaction, the setting and the cultural background [6]. Given the positive attitudes and expectation towards complementary therapies it seems possible that complex rituals such as acupuncture could provoke significant psychobiological responses.

The most straightforward way to investigate whether sham acupuncture is associated with larger effects than a pharmacological placebo would be in randomized trials including both these interventions. The only trial using such an approach indeed found a significant superiority of sham acupuncture [18]. Another, albeit methodologically weaker, possibility is to compare differences between sham acupuncture interventions and no treatment control groups in acupuncture trials with those of (other) placebos and no treatment control groups in other trials. Hrobjartsson and Gøtzsche have repeatedly reviewed all available trials including both a placebo or sham and a no treatment group for any condition [19-21]. The latest

update of their Cochrane review includes a total of 234 trials. In a pre-planned subgroup analysis they found that studies using “physical placebos” (including sham acupuncture) reported larger placebo effects (standardized mean difference (SMD) -0.31; 95%-confidence interval (CI) -0.41, -0.22) than studies using “pharmacological placebos” (SMD -0.10; 95%CI -0.20, -0.01) [21]. In a re-analysis of their data we separated the trials in which the physical placebo was sham acupuncture from those which used other physical placebos. Effects sizes were significantly larger in trials using sham acupuncture than in trials using other physical placebos (SMDs -0.41 (-0.56, -0.24) vs. -0.26 (-0.37, -0.15); p = 0.007) [22].
The Cochrane review and our re-analysis of these data did not include a number of recent rigorous, large acupuncture trials which include both a sham and a no treatment group. Furthermore, these reviews did not investigate whether large non-specific effects might make it difficult to detect specific effects. Therefore, we have performed a systematic review of acupuncture trials in any condition including both a sham and a no treatment group published until April 2010. Our primary aim was to investigate the size of non-specific effects of acupuncture (difference sham acupuncture vs. no acupuncture). Our secondary aims were to investigate factors (such as type of sham intervention, condition, study quality, or intensity of co-interventions) possibly influencing the size of such non-specific effects and to quantify specific (difference acupuncture vs. sham acupuncture) and total effects of acupuncture (difference acupuncture vs. no acupuncture) in the included trials.

Methods

Selection criteria

To be included studies had to meet the following criteria: (1) design – trials in which allocation to groups was explicitly randomized; (2) participants – persons treated for any illness or for preventative purposes. Trials in healthy volunteers measuring physiological outcomes were excluded; (3) intervention – insertion of needles described as acupuncture at acupuncture points, pain or trigger points with or without stimulation. Trials on interventions without skin penetration (e.g. laser acupuncture) were excluded; (4) sham intervention –

interventions described as sham, placebo, dummy or fake treatment which differed from true acupuncture at least in one of two key aspects (skin penetration or point location); (5) no acupuncture control group – there had to be a second control group in which participants received neither true nor sham acupuncture. Participants could be either completely untreated or receive treatments which were also received in the true and sham acupuncture groups (e.g. rescue medication, basic treatment or routine care); (6) outcomes – any clinical outcomes for which the calculation of an effect size estimate was considered possible.

Data sources and searches

For identifying potentially relevant studies we searched Medline (from 1966 to April 2010), Embase (from 1988 to April 2010) for all sham-controlled trials of acupuncture (see appendix for detailed search strategies): Furthermore, we searched the Cochrane Central Register of Controlled Trials using a search strategy based on a Cochrane review of randomized trials with placebo and no treatment controls in all medicine [20]. While Chinese trials identified by our search were eligible, we did not specific Chinese databases. One reviewer screened titles and abstracts of all references identified and excluded those which were clearly irrelevant. Full texts of all remaining articles were obtained and assessed independently for eligibility by two reviewers. Disagreements or uncertainties were resolved by discussion.

Data extraction and quality assessment

One reviewer extracted information on the following aspects from included studies using a standard form: diagnosis, recruitment, number and type of study centres, number and types of intervention and control groups, details on acupuncture and sham interventions, how patients were informed about these, qualification of acupuncturists, co-interventions, study duration, number of patients randomized, analyzed and dropping out (per group), age, gender, results on the main outcome measures, important secondary outcomes and responder data. A second reviewer checked all extractions of study results against the original publications. Trials were considered to have lower risk of bias if they reported an

adequate method of randomization concealment and had a drop-out rate below 15% [23]. For our main analyses we used the following strategy to choose the outcome: 1) it should be a continuous outcome (mean and standard deviation available, or the standard deviation could be calculated from standard errors, confidence intervals etc.; we did not impute standard deviations for studies without available data on variability or precision); 2) the timing should be as close as possible to the completion of treatment; 3) when there was a clearly predefined main outcome measure we chose this measure (but always preferred the measurement at the end of treatment over other time points or change from baseline); 4) when there was no predefined single main outcome measure two reviewers independently chose the outcome considered most important (two disagreements were resolved by discussion). 5) If available, we used intention to treat data; otherwise we used the data as presented in the publication. If a trial had more than one intervention (for example, an individualized and a standardized intervention) or sham group the data was pooled. For more recent studies we tried to contact authors inquiring for further information if data for meta-analysis were missing.

Data synthesis and analysis

The Cochrane Collaboration's Review Manager (RevMan) 5 software was used for meta-analyses. Three comparisons were investigated: sham acupuncture vs. no acupuncture (primary comparison), acupuncture vs. sham acupuncture, and acupuncture vs. no acupuncture. Studies were categorized into the clinical categories chronic pain studies, short term studies (studies with an observation period of less than three days) and other studies. The main analysis was based on trials reporting a continuous outcome measure using the standardized mean difference (SMDs; difference between the means/pooled standard deviation) as effect size estimate. As we assumed that studies would be clinically heterogeneous a random effects model with the inverse variance method was used for meta-analysis. Negative SMDs indicate a beneficial effect of sham acupuncture over no acupuncture, acupuncture over sham acupuncture and acupuncture over no acupuncture,

respectively. SMDs ≤ -0.4 were considered small effects, those between -0.41 and -0.7 moderate and those > -0.7 large effects [24]. To investigate statistical heterogeneity RevMan 5 uses Tau^2, Chi^2 and I^2. We considered I^2-values between 30% and 60% as indicating moderate heterogeneity and higher values as indicating substantial heterogeneity. Subgroup comparisons were performed using the method by Deeks et al. [25] implemented in RevMan 5. Egger's test was used to assess funnel plot asymmetry [26].

To check the robustness of results we performed sensitivity analyses a) including three-armed studies which had been excluded as they did not meet all inclusion criteria but still could be considered addressing the questions investigated in this review ("borderline" studies; see results); b) using different outcomes for studies with more than one relevant outcome at completion of treatment; c) using dichotomous outcome measures (with a relative risk < 1 indicating a beneficial effect).

For exploratory analyses we defined further subgroups: larger (at least 100 patients) and smaller (< 100 patients) comparisons, lower and higher risk of bias (see above); studies with intense or less intense co-interventions in all study arms; with and without skin penetration (and depending on where needles were placed) in sham groups; studies with and without a clearly defined main outcome measure; and studies describing sham in the consent procedure as another treatment or placebo. In multivariate random effects meta-regression analyses we investigated simultaneously the influence of risk of bias, co-interventions, skin penetration in the sham group and condition (chronic pain vs. others). Analyses were carried out using the restricted information maximum likelihood (REML) method. For meta-regression analyses the PASW (SPSS; Chicago, Illinois) v17.0 and 18.0 software using additional macros by Wilson was used [27]. To investigate the hypothesis that there is an inverse correlation between specific and non-specific effects (trials with large non-specific effects are less likely to find large specific effects than trials with small non-specific effects) we performed a linear regression analysis using the inverse of the squared pooled standard error as weighting factor.

Results

Literature search and selection

The literature search identified a total of 1854 references of which 1779 were excluded in the screening process as they clearly did not meet inclusion criteria (see figure 1). The full text of the remaining 75 references was formally assessed for eligibility. A total of 37 studies [28-64] met the inclusion criteria. Eleven additional publications reported protocols or treatment details of trials included in the review or reported the same results in another language (see web appendix table 1). Eighteen articles did not meet inclusion criteria and two were protocols of ongoing trials (see web appendix table 2). Two abstracts reported minimal information on probably eligible trials including results for a dichotomous outcome [65-66]; attempts to obtain further information from the authors were unsuccessful. In four further studies patients in the no acupuncture control group received minor interventions not provided in the other two groups [67-70]. Finally, for one study presenting an asymmetric confidence interval for the continuous main outcome measure we were unable to unambiguously calculate the standard deviation [71]. The latter five trials were included in a sensitivity analysis as “borderline” studies.

Description of included studies

The 37 eligible trials included a total of 5754 patients (median 75, minimum 30, maximum 638). Fourteen trials (3369 patients) addressed chronic pain or a condition associated with chronic pain (table 1); eight were short-term trials with duration of less than three days (522 patients; table 2) investigating whether acupuncture is helpful for sedation, anxiety, pain or nausea associated with surgical operations, endoscopic interventions or labour; and 15 trials (1863 patients) addressed a variety of other conditions (table 3). Ten of the 14 chronic pain trials, but only six of the remaining 23 studies, reported an adequate method of allocation concealment. Drop-out rates were between 54% and 95% in three addiction trials but low in most other studies. Ten chronic pain trials and three trials on other conditions both reported

an adequate method of allocation concealment and a drop-out rate below 15%, and were classified as having a lower risk of bias.

Fifteen studies had a clearly predefined main outcome measure. For 32 trials a continuous effect size measure could be calculated, for 24 trials a dichotomous (for 19 trials both). Acupuncture interventions varied strongly regarding number of sessions, type of acupuncture (classical acupuncture, electro-acupuncture, ear acupuncture), level of individualisation for point selection, and number of needles used. In 31 trials the sham procedure involved skin penetration (in seven trials at acupuncture points not indicated for the condition treated and in 24 trials outside known acupuncture points); six trials used approaches without skin penetration (in three trials at the same points as in the acupuncture group and in three trials outside known points).

Meta-analysis of non-specific effects (sham acupuncture vs. no acupuncture)

The main analyses are based on the 32 trials reporting data on a continuous outcome. For the comparison of sham acupuncture with no acupuncture, the pooled standardized mean differences (SMD) were -0.53 (95%CI -0.67, -0.39) among chronic pain trials, -0.23 (-0.50, 0.04) among short-term studies and -0.42 (95%CI -0.66, -0.18) in other studies (figure 2). The test for differences between diagnostic subgroups missed statistical significance at the 5% level ($p = 0.08$). Effect sizes showed moderate statistical heterogeneity among chronic pain studies, no heterogeneity among short term studies and marked heterogeneity among the other studies. If studies were pooled across clinical subgroups the SMD was -0.45 (95%CI -0.57, -0.34). In seven trials effects over no treatment groups were large (SMDs were above -0.7), in nine trials moderate (between -0.4 and -0.7), and in 16 small (< -0.4). Results were similar when borderline studies were included, when in studies without a predefined main outcome measure other outcomes were chosen or when dichotomous outcomes were analyzed (see web appendix table 3). Egger's test did not suggest funnel plot asymmetry ($p = 0.25$; asymmetry coefficient 0.21, figure 3). In exploratory subgroup analyses (see web appendix table 3) effect sizes differed significantly according to the level of co-intervention

(larger if less co-interventions) and according to the type of sham intervention (larger if no skin penetration). Non-specific effects tended to be larger in trials with larger sample size, lower risk of bias, and a clearly predefined outcome, but the differences were not statistically significant. In multivariate meta-regression analyses only the association with level of co-interventions approached statistical significance ($p = 0.07$). Trials with larger effects of sham over no acupuncture reported smaller effects of acupuncture over sham intervention than trials with smaller non-specific effects ($\beta = -0.39$, $p = 0.029$).

Meta-analysis of specific effects (acupuncture vs. sham acupuncture) and total effects (acupuncture vs. no acupuncture)

For the comparison of acupuncture with sham acupuncture, the pooled random effects SMDs were -0.46 (95%CI -0.72, -0.20) for chronic pain studies, -0.34 (95%CI -0.79, 0.12) for short term studies, and -0.28 (-0.59, 0.03) for other studies (see web appendix figure 1). There were no statistically significant ($p = 0.71$) differences between diagnostic subgroups, but there was substantial statistical heterogeneity in all three clinical categories. If trials were pooled across categories the SMD was -0.37 (95%CI -0.55, -0.19). The funnel plot was highly asymmetrical (web appendix figure 2; $p = 0.002$; asymmetry coefficient -0.52). Larger trials yielded significantly less positive results than smaller trials (SMDs -0.15 (95%CI -0.31, 0.01) vs. -0.59 (95%CI -0.93, -0.24), $p < 0.001$). Specific effects were also smaller in trials with lower risk of bias and more intense co-interventions, while skin penetration and condition did not have a significant influence.

The pooled SMDs between acupuncture and no acupuncture were -0.94 (95%CI -1.20, -0.67) for chronic pain studies, -0.60 (95%CI -1.08, -0.12) for short term studies, and -0.63 (-0.91, -0.35) for other studies (see web appendix figure 3) with marked heterogeneity in all three categories. If all studies were pooled the SMD was -0.77 (95%CI -0.94, -0.59). There was significant ($p = 0.03$; asymmetry coefficient -0.38) funnel plot asymmetry with smaller studies yielding larger effect estimates (see web appendix figure 4 for the funnel plot).

Discussion

Summary of main findings

According to our findings (sham) acupuncture interventions are often associated with noteworthy non-specific effects. Differences between sham acupuncture and no acupuncture groups tended to be smaller in trials in which there were intense co-interventions in all study groups. Indicators of study quality (sample size, risk of bias, predefinition of a main outcome measure) were not associated significantly with effect size. Trials with larger effects of sham over no acupuncture reported smaller effects of acupuncture over sham intervention than trials with smaller non-specific effects. In our analyses we also found small to moderate specific effects of acupuncture interventions over sham acupuncture, however, trials with large sample size and low risk of bias yielded less positive results. In our study set the total effect of acupuncture interventions including both specific and non-specific effects was, on average, at least moderate in size.

Strengths and limitations

Although we did not systematically search Chinese language databases our review is currently the most comprehensive and largest analysis of randomized trials of acupuncture including both a sham and a no treatment control group. It includes many more and larger trials than previous analyses [21-23]. The overall findings are highly robust to sensitivity analyses and indicators of study quality. The most important limitation of our review is the strong heterogeneity of our trial set regarding patients, interventions, outcomes and methodological quality. We do not think that pooling such a heterogeneous set of studies would be adequate if the aim was primarily to assess effectiveness for clinical decision-making. However, our primary aim was to investigate whether (sham) acupuncture interventions are, on average, associated with relevant non-specific effects. To assess the size of non-specific effects it is necessary to include trials with both a sham and a no acupuncture control group. As the number of such trials is limited pooling all available

information can be justified for generating hypotheses and has been performed in the Cochrane review on placebo effects in all conditions in a much more radical manner [21]. The comparisons between sham acupuncture and acupuncture in the primary studies included in our review are unblinded. As almost all trials focussed on patient-reported outcome measures there is considerable risk of bias. Patients randomized to the no treatment group might be disappointed and experience “nocebo” effects, or might give overly negative ratings for subjective symptoms. On the other hand, patients randomized to no treatment groups might use larger doses of rescue medication or co-interventions which would lead to an underestimation of the differences. In fact, in some of the trials included in our review patients in no acupuncture control groups had higher analgesic use than patients in the sham groups (e.g. [51, 53]). Insufficient blinding is also a problem for the comparison between acupuncture and placebo acupuncture [23]. However, if patients find out that they are in a sham group one would expect rather an underestimation of the effect of sham over no treatment. In summary, it is difficult to assess to what extent and in which direction biases can distort effect estimates between sham and no acupuncture groups. It is noteworthy that although indicators of study quality were not significantly associated with the size of non-specific effects, better and larger studies tended to report larger effects. It seems that our estimate of non-specific effects is less subject to small study bias and other biases than those for specific and total effects.

Interpretation

Our findings are highly consistent with smaller analyses available in the literature [22-23]. The re-analysis of the 21 acupuncture trials included in the Cochrane review on placebo effects yielded a SMD of -0.41 [22]. Due to slightly different inclusion criteria five trials were excluded for the current analyses. A meta-analysis by Madsen et al. [23] who reviewed 13 three-armed trials on acupuncture for acute and chronic pain found a SMD of -0.42. Nine of the studies included in their review were also included by us while we excluded four trials

due to slightly different selection criteria. Our main analysis includes 23 additional trials (including seven trials addressing chronic or acute pain).

It has been argued that sham interventions in which needles penetrate the skin (particularly if applied in the same dermatomes as the true acupuncture intervention) are physiologically not inert and, therefore, should not be considered as placebos [5]. Our exploratory subgroup analyses (as well as similar analyses in the review by Madsen et al. [23]) do not provide evidence that sham interventions involving needle penetration are associated with larger non-specific than those which do not. Thus, the limited available data suggests that skin penetration or no skin penetration does not seem to make a big difference.

If acupuncture should have indeed relevant total effects but only very limited specific effects this would have major implications for the conduct and interpretation of clinical trials. Based on our data and available systematic reviews [2-5, 72] it seems reasonable to assume an average SMD of 0.4 (or more) for non-specific effects and of 0.2 SMD (or less) for specific effects at least for a number of conditions. To achieve 80% power a two-armed sham-controlled clinical trial investigating a specific effect of 0.2 SMD would have to recruit about 800 patients. This suggests that almost all available trials comparing true and sham acupuncture would be underpowered.

One could argue that a SMD of 0.2 is clinically irrelevant. In line with that reasoning Madsen et al. questioned in their review “the prevailing hypothesis that acupuncture has an important effect on pain in general” [23]. However, we believe that another conclusion is possible, too. As ourselves, Madsen et al. have found, on average, a moderately large effect of sham interventions over no acupuncture groups, and both reviews found at least small specific effects of acupuncture over sham interventions. The total effect of acupuncture seems to be at least moderate in size in a number of conditions, and such effects can well be clinically relevant. For many established drug treatments SMDs over placebo are in the range between 0.3 and 0.5 (e.g. [75-76]). If, as the available data suggests [21] clinical effects associated with pharmacological placebos are small compared to no treatment (with a SMD of 0.1 on average) the total effects of these treatments could be in a similar range (around a

SMD of 0.4 to 0.6) as those of several acupuncture interventions. It could be argued that for a suffering individual it does not matter whether relief is due to specific or non-specific effects. However, as the evidence for larger non-specific effects of acupuncture compared to other treatments comes with one exception [18] from indirect comparisons open to confounding firm conclusions are not yet possible.

We think that our findings are of major relevance to the question of how the clinical effectiveness of complex non-drug interventions should be assessed. It is likely that non-specific effects vary between different types of complex treatment interventions. The concept of specific and non-specific effects might not be fully adequate in that case, as so-called non-specific effects might turn out to be characteristic for a given therapeutic setting. If the total effect of an intervention in clinical practice would indeed consist of variable contributions of specific and non-specific effects it could be that a treatment which has only minor or even no specific but clinically relevant non-specific effects has a larger total effect than a treatment with moderate specific but only minor non-specific effects. This has been denoted the efficacy paradox [72]. Should such a treatment be readily available? The position of a pragmatic decision-maker could be yes if the comparator treatment represents adequate standard treatment. In fact, in Germany acupuncture is routinely reimbursed for chronic low back pain as in a large randomized trial acupuncture (but also sham acupuncture which is not reimbursed) was more effective than treatment based on German guidelines [73]. Sceptical scientists would argue that these results are likely to be biased due to lack of blinding and that acupuncture should not be considered effective. Furthermore, if issues such as expectancies, beliefs and trust should have a relevant influence on the effectiveness of a treatment, findings of clinical trials might no longer be valid when attitudes in a population change over time.

Conclusions

Sham acupuncture interventions are often associated with moderately large non-specific effects which could make it difficult to detect small additional specific effects. Compared to

inert placebo interventions effects associated with sham acupuncture might be larger which would have considerable implications for the design and interpretation of clinical trials. Total effects of acupuncture interventions including both specific and non-specific effects often seem to be at least moderate in size. We believe that there has to be a discussion involving scientists, decision-makers, health care providers and patients whether and when the evidence for clinically relevant total effects from non-blinded comparisons is sufficient to consider a treatment effective even if specific effects due to the postulated mechanism of action might be minor or even non-existing.

Competing interests

Klaus Linde has received travel reimbursement and fees for speaking at conferences organized by acupuncture societies in the USA, UK, Germany, Japan and Spain. Antonius Schneider has received fees for lecturing for a German acupuncture society (DÄGfA) until 2006. Karin Meissner and Karin Niemann do not have any conflicts of interest.

Authors' contributions

KL, KN and KM were involved in the literature search, data extraction and analysis. AS provided advice on acupuncture and participated in the interpretation of the data. KL conceived and coordinated the study and wrote the first draft of the manuscript. All authors commented on drafts and approved the final manuscript.

Acknowledgements

Funding for this project was neither applied for nor received.

The authors would like to thank Asbjørn Hrobjartsson, Levente Kriston, Mike Cummings and Benno Brinkhaus for critically commenting on earlier versions of the manuscript. Part of this

work has been performed for the MD thesis of Karin Niemann at the Technische Universität München.

References

1. Haake M, Muller HH, Schade-Brittinger C, Basler HD, Schafer H, Maier C, Endres HG, Trampisch HJ, Molsberger A: **German Acupuncture Trials (GERAC) for chronic low back pain: randomized, multicenter, blinded, parallel-group trial with 3 groups**. *Arch Intern Med* 2007, **167**:1892-1898.
2. Linde K, Allais G, Brinkhaus B, Manheimer E, Vickers A, White AR: **Acupuncture for migraine prophylaxis**. *Cochrane Database Syst Rev* 2009(1):CD001218.
3. Bausell RB: **Snake oil science. The truth about complementary and alternative medicine**. Oxford: Oxford University Press; 2007.
4. 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 Alternat Complement Med* 2006, **12**:303-310.
5. Lund I, Lundeberg T: **Are minimal, superficial or sham acupuncture procedures acceptable as inert placebo controls?** *Acupunct Med* 2006, **24**:13-15.
6. Kaptchuk TJ: **The placebo effect in alternative medicine: Can the performance of a healing ritual have clinical significance?** *Ann Intern Med* 2002, **136**:817-825.
7. Liu T, Yu CP: **Placebo analgesia, acupuncture and sham surgery**. *eCAM* 2010 (epub ahead of print doi:10.1093/ecam/neq030).
8. Shapiro AK, Morris LA: **The placebo effect in medical and psychological therapies**. In: *Handbook of psychotherapy and behavior change.* Edited by Garfield SL, Bergin AE, vol. 2nd edition. New York: Wiley; 1978: 369-410.

9. Grünbaum A: **The placebo concept in medicine and psychiatry**. *Psycholog Med* 1986, **16**:19-38.

10. Finniss DG, Kaptchuk TJ, Miller F, Benedetti F: **Biological, clinical, and ethical advances of placebo effects**. *Lancet* 2010, **375**:686-695.

11. Ernst E, Resch KL: **Concept of true and perceived placebo effects**. *BMJ* 1995, **311**:551-553.

12. Kienle GS, Kiene H: **The powerful placebo effect: fact or fiction?** *J Clin Epidemiol* 1997, **50**:1311-1318.

13. Vickers AJ, de Craen AJM: **Why use placebos in clinical trials? A narrative review of the methological literature**. *J Clin Epidemiol* 2000, **53**:157-161.

14. Wampold BE, Minami T, Tierney SC, Baskin TW, Bhati KS: **The placebo is powerful: estimating placebo effects in medicine and psychotherapy from randomized clinical trials**. *J Clin Psychol* 2005, **61**:835-854.

15. Napadow V, Ahn A, Longhurst J, Lao L, Stener-Victorin E, Harris R, Langevin HM: **The status and future of acupuncture mechanism research**. *J Altern Complement Med* 2008, **14**:861-869.

16. Dincer F, Linde K: **Sham interventions in randomized clinical trials of acupuncture - a review**. *Complement Ther Med* 2003, **11**:235-242.

17. Kaptchuk TJ, Kelley JM, Conboy LA, Davis RB, Kerr CE, Jacobson EE, Kirsch I, Schyner RN, Nam BH, Nguyen LT, Park M, Rivers AL, McManus C, Kokkotou E, Drossman DA, Goldman P, Lembo AJ: **Components of placebo effect: randomised controlled trial in patients with irritable bowel syndrome**. *BMJ* 2008, **336**:999-1003.

18. Kaptchuk TJ, Stason WB, Davis RB, Legedza ATR, Schnyer RN, Kerr CE, Stone DA, Huyn Nam B, Kirsch I, Goldman RH: **Sham device vs. inert pill: randomised controlled trial of two placebo treatments**. *BMJ* 2006, **332**:391-397.

19. Hrobjartsson A, G¢tzsche PC: **Is the placebo powerless? An analysis of clinical trials comparing placebo with no treatment**. *N Engl J Med* 2001, **344**:1594-1602.

20. Hrobjartsson A, G¢tzsche PC: **Placebo interventions for all clinical conditions**. *Cochrane Database Syst Rev* 2004(3):CD003974.

21. Hrobjartsson A, G¢tzsche PC: **Placebo interventions for all clinical conditions**. *Cochrane Database Syst Rev* 2010(1):CD003974.

22. Linde K, Niemann K, Meissner K: **Are sham acupuncture interventions more effective than (other) placebos? A re-analysis of data from the Cochrane review on placebo effects**. *Forsch Komplementrmed* 2010, **17** (epub ahead of print, doi: 10.1159/000320374).

23. Madsen MV, G¢tzsche PC, Hrobjartsson A: **Acupuncture treatment for pain: systematic review of randomised clinical trials with acupuncture, placebo acupuncture, and no acupuncture groups**. *BMJ* 2009, **338**:a3115.

24. Schünemann HJ, Oxman AD, Vist GE, Higgins JPT, Deeks JJ, Glasziou P, Guyatt GH: **Interpreting results and drawing conclusions**. In: *Cochrane Handbook for Systematic Reviews of Interventions.* Edited by Higgins JPT, Green S: The Cochrane Collaboration; 2008. Available at http://www.cochrane-handbook.org/

25. Deeks JJ, Altman DG, Bradburn MJ: **Statistical methods for examining heterogeneity and combining results from several studies in meta-analysis**. In: *Sytematic reviews in health care - meta-analysis in context.* Edited by Egger M, Smith GD, Altman DG. London: BMJ Books; 2001: 285-312.

26. Egger M, Davey SG, Schneider M, Minder C: **Bias in meta-analysis detected by a simple, graphical test**. *BMJ* 1997, **315**:629-634.

27. Wilson DB: **SPSS, Stata, and SAS macros for performing meta-analytic analyses**. 2010. Avaliable at http://mason.gmu.edu/~dwilsonb/ma.html

28. Allen JJB, Schnyer R, Hitt SK: **The efficacy of acupuncture in the treatment of major depression in women**. *Psychol Sci* 1998, **9**:397-401.

29. Allen JJB, Schnyer RN, Chambers AS, Hitt SK, Moreno FA, Manber R: **Acupuncture for depression: a randomized controlled trial**. *J Clin Psychiatry* 2006, **67**:1665-1673.

30. Asher GN, Coeytaux RR, Chen W, Reilly AC, Loh YL, Harper TC: **Acupuncture to initiate labor (Acumoms 2): a randomized, sham-controlled clinical trial**. *J Matern Fetal Neonatal Med* 2009, **22**:843-848.

31. Aune A, Alraek T, LiHua H, Baerheim A: **Acupuncture in the prophylaxis of recurrent lower urinary tract infection in adult women**. *Scand J Prim Health Care* 1998, **16**:37-39.

32. Avis NE, Legault C, Coeytaux RR, Pian-Smith M, Shifren JL, Chen W, Valaskatgis P: **A randomized, controlled pilot study of acupuncture treatment for menopausal hot flashes**. *Menopause* 2008, **15**:1070-1078.

33. Birch S, Jamison RN: **Controlled trial of Japanese acupuncture for chronic myofascial neck pain: assessment of specific and nonspecific effects of treatment**. *Clin J Pain* 1998, **14**:248-255.

34. Brinkhaus B, Witt CM, Jena S, Linde K, Streng A, Wagenpfeil S, Irnich D, Walther HU, Melchart D, Willich SN: **Acupuncture in patients with chronic low back pain: a randomized controlled trial**. *Arch Intern Med* 2006, **166**:450-457.

35. Bullock ML, Kiresuk TJ, Pheley AM, Culliton PD, Lenz SK: **Auricular acupuncture in the treatment of cocaine abuse. A study of efficacy and dosing**. *J Subst Abuse Treat* 1999, **16**:31-38.

36. Cabrini L, Gioia L, Gemma M, Melloni G, Carretta A, Ciriaco P, Puglisi A: **Acupuncture for diagnostic fiberoptic bronchoscopy: a prospective, randomized, placebo-controlled study**. *Am J Chin Med* 2006, **34**:409-415.

37. Cherkin DC, Sherman KJ, Avins AL, Erro JH, Ichikawa L, Barlow WE, Delaney K, Hawkes R, Hamilton L, Pressman A, Khalsa PS, Deyo RA: **A randomized trial comparing acupuncture, simulated acupuncture, and usual care for chronic low back pain**. *Arch Intern Med* 2009, **169**:858-866.

38. Dundee JW, Chestnutt WN, Ghaly RG, Lynas AG: **Traditional Chinese acupuncture: a potentially useful antiemetic?** *Br Med J (Clin Res Ed)* 1986, **293**:583-584.

39. Facco E, Liguori A, Petti F, Zanette G, Coluzzi F, De Nardin M, Mattia C: **Traditional acupuncture in migraine: a controlled, randomized study**. *Headache* 2008, **48**:398-407.

40. Fanti L, Gemma M, Passaretti S, Guslandi M, Testoni PA, Casati A, Torri G: **Electroacupuncture analgesia for colonoscopy. a prospective, randomized, placebo-controlled study**. *Am J Gastroenterol* 2003, **98**:312-316.

41. Foster NE, Thomas E, Barlas P, Hill JC, Young J, Mason E, Hay EM: **Acupuncture as an adjunct to exercise based physiotherapy for osteoarthritis of the knee: randomised controlled trial**. *BMJ* 2007, **335**:436.

42. Freire AO, Sugai GCM, Chrispin FS, Togeiro SM, Yamamura Y, Mello LE, Tufik S: **Treatment of moderate obstructive sleep apnea syndrome with acupuncture: a randomised, placebo-controlled pilot trial**. *Sleep Med* 2007, **8**:43-50.

43. Gioia L, Cabrini L, Gemma M, Fiori R, Fasce F, Bolognesi G, Spinelli A, Beretta L: **Sedative effect of acupuncture during cataract surgery: prospective randomized double-blind study**. *J Cataract Refract Surg* 2006, **32**:1951-1954.

44. Helms JM: **Acupuncture for the management of primary dysmenorrhea**. *Obstet Gynecol* 1987, **69**:51-56.

45. Karst M, Winterhalter M, Munte S, Francki B, Hondronikos A, Eckardt A, Hoy L, Buhck H, Bernateck M, Fink M: **Auricular acupuncture for dental anxiety: a randomized controlled trial**. *Anesth Analg* 2007, **104**:295-300.

46. Kotani N, Kushikata T, Suzuki A, Hashimoto H, Muraoka M, Matsuki A: **Insertion of intradermal needles into painful points provides analgesia for intractable abdominal scar pain**. *Reg Anesth Pain Med* 2001, **26**:532-538.

47. Lee SH, Lee BC: **Electroacupuncture relieves pain in men with chronic prostatitis/chronic pelvic pain syndrome: three-arm randomized trial**. *Urology* 2009, **73**:1036-1041.

48. Leibing E, Leonhardt U, Koster G, Goerlitz A, Rosenfeldt JA, Hilgers R, Ramadori G: **Acupuncture treatment of chronic low-back pain - a randomized, blinded, placebo-controlled trial with 9-month follow-up**. *Pain* 2002, **96**:189-196.

49. Lembo AJ, Conboy L, Kelley JM, Schnyer RS, McManus CA, Quilty MT, Kerr CE, Drossman D, Jacobson EE, Davis RB: **A treatment trial of acupuncture in IBS patients**. *Am J Gastroenterol* 2009, **104**:1489-1497.

50. Li CK, Nauck M, Loser C, Folsch UR, Creutzfeldt W: **[Acupuncture to alleviate pain during colonoscopy]**. *Dtsch Med Wochenschr* 1991, **116**:367-370.

51. Linde K, Streng A, Jurgens S, Hoppe A, Brinkhaus B, Witt C, Wagenpfeil S, Pfaffenrath V, Hammes MG, Weidenhammer W, Willich SN, Melchart D: **Acupuncture for patients with migraine: a randomized controlled trial**. *JAMA* 2005, **293**:2118-2125.

52. Medici TC, Grebski E, Wu J, Hinz G, Wuthrich B: **Acupuncture and bronchial asthma: a long-term randomized study of the effects of real versus sham acupuncture compared to controls in patients with bronchial asthma**. *J Altern Complement Med* 2002, **8**:737-750.

53. Melchart D, Streng A, Hoppe A, Brinkhaus B, Witt C, Wagenpfeil S, Pfaffenrath V, Hammes M, Hummelsberger J, Irnich D, Weidenhammer W, Willich SN, Linde K: **Acupuncture in patients with tension-type headache: randomised controlled trial**. *BMJ* 2005, **331**:376-382.

54. Molsberger AF, Mau J, Pawelec DB, Winkler J: **Does acupuncture improve the orthopedic management of chronic low back pain--a randomized, blinded, controlled trial with 3 months follow up**. *Pain* 2002, **99**:579-587.

55. Rampes H, Pereira S, Mortimer A, Manoharan S, Knowles M: **Does electroacupuncture reduce craving for alcohol? A randomized controlled study**. *Complement Ther Med* 1997, **5**:19-26.

56. Röschke J, Wolf C, Muller MJ, Wagner P, Mann K, Grozinger M, Bech S: **The benefit from whole body acupuncture in major depression**. *J Affect Disord* 2000, **57**:73-81.

57. Rusy LM, Hoffman GM, Weisman SJ: **Electroacupuncture prophylaxis of postoperative nausea and vomiting following pediatric tonsillectomy with or without adenoidectomy**. *Anesthesiology* 2002, **96**:300-305.

58. Smith C, Crowther C, Beilby J: **Acupuncture to treat nausea and vomiting in early pregnancy: a randomized controlled trial**. *Birth* 2002, **29**:1-9.

59. Suarez-Almazor ME, Looney C, Liu Y, Cox V, Pietz K, Donald M, Street R: **A randomized controlled trial of acupuncture for osteoarthritis of the knee: effects of patient-provider communication**. *Arthritis Care Res* 2010; **62**:1229-1236.

60. Tremeau ML, Fontanie-Ravier P, Teurnier F, Demouzon J: **[Protocol of cervical maturation by acupuncture]**. *J Gynecol Obstet Biol Reprod (Paris)* 1992, **21**:375-380.

61. Wang S-M, Dezinno P, Lin EC, Lin H, Yue JJ, Berman MR, Braveman F, Kain ZN: **Auricular acupuncture as a treatment for pregnant women who have low back**

and posterior pelvic pain: a pilot study. *Am J Obstet Gynecol* 2009, **201**:271.e271-279.

62. Witt C, Brinkhaus B, Jena S, Linde K, Streng A, Wagenpfeil S, Hummelsberger J, Walther HU, Melchart D, Willich SN: **Acupuncture in patients with osteoarthritis of the knee: a randomised trial**. *Lancet* 2005, **366**:136-143.

63. Worner TM, Zeller B, Schwarz H, Zwas F, Lyon D: **Acupuncture fails to improve treatment outcome in alcoholics**. *Drug Alcohol Depend* 1992, **30**:169-173.

64. Ziaei S, Hayipour L: **Effect of acupuncture on labor**. *Intern J Gynecol Obstet* 2006, **92**:71-72.

65. Benson MR, Elkind-Hirsch KE, Theall A, Fong K, Hogan RB, Scott RT: **Impact of acupuncture before and after embryo transfer on the outcome of in vitro fertilization cycles: a prospective single-blind randomized study**. *Fertil Steril* 2006, **86 (suppl 1)**:S135.

66. Fratterelli JL, Leondires MR, Fong K, Theall A, Locatelli S, Scott RT: **Laser acupuncture before and after embryo transfer imporves ART delivery rates: results of a prospective randomized double-blinded placebo controlled five-armed trial involving 1000 patients**. *Fertil Steril* 2008, **90 (suppl 1)**:S105.

67. Avants SK, Margolin A, Holford TR, Kosten TR: **A randomized controlled trial of auricular acupuncture for cocaine dependence**. *Arch Intern Med* 2000, **160**:2305-2312.

68. Berman BM, Lao L, Langenberg P, Lee WL, Gilpin AMK, Hochberg MC: **Effectiveness of acupuncture as adjunctive therapy in osteoarthritis of the knee: a randomized, controlled trial.** *Ann Intern Med* 2004, **141**:901-910.

69. Margolin A, Kleber HD, Avants SK, Konefal J, Gawin F, Stark E, Sorensen J, Midkiff E, Wells E, Jackson TR, Bullock M, Culliton PD, Boles S, Vaughan R: **Acupuncture**

for the treatment of cocaine addiction: a randomized controlled trial. *JAMA* 2002, **287**:55-63.

70. Scharf H-P, Mansmann U, Streitberger K, Witte S, Kramer J, Maier C, Trampisch H-J, Victor N: **Acupuncture and knee osteoarthritis: a three-armed randomized trial.** *Ann Intern Med* 2006, **145**:12-20.

71. Shen J, Wenger N, Glaspy J, Hays RD, Albert PS, Choi C, Shekelle PG: **Electroacupuncture for control of myeloablative chemotherapy-induced emesis: A randomized controlled trial**. *JAMA* 2000, **284**:2755-2761.

72. Walach H: **The efficacy paradox in randomized controlled trials of CAM and elsewhere: beware of the placebo trap**. *J Altern Complement Med* 2001, **7**:213-218.

73. Haake M, Muller HH, Schade-Brittinger C, Basler HD, Schafer H, Maier C, Endres HG, Trampisch HJ, Molsberger A: **German Acupuncture Trials (GERAC) for chronic low back pain: Randomized, multicenter, blinded, parallel-group trial with 3 groups**. *Arch Intern Med* 2007, **167**:1892-1898.

Figures

Figure 1 - Flow chart

Figure 2 - The non-specific effect of acupuncture (difference between groups receiving sham acupuncture and no acupuncture)

SD = standard deviation; Total = number of patients; 95% CI = 95% confidence interval; IV = inverse variance method; Random = random effects model; df = degrees of freedom

Figure 3 - Funnel plot of studies comparing sham acupuncture vs. no acupuncture

SE = standard error; SMD = standardized mean difference

Table 1 - Characteristics of included trials: Chronic pain trials

Trial	Clinical problem	Sample size (% drop-out	Conceal-ment	Outcome used for meta-analysis	Intervention details		Standard basic care in all groups
Birch 1998 [33]	Chonic myofascial neck pain	46 (22%)	Unclear	Change from baseline on pain intensity rating scale	N: 14 D: 30 T: 12w	Ac: C JA S: I C 2	NSAIDs if needed
Brinkhaus 2006 [34]	Chronic low back pain	298 (6%)	Adequate	PMOM: VAS pain intensity week 8	N: 12 D: 30 T: 8w	Ac: B CA S: I B 1	NSAIDs if needed
Cherkin 2009 [37]	Chronic low back pain	638 (5%)	Adequate	PMOM: Roland Morris Disability Quest. week 8	N: 10 D: 20 T: 7w	Ac1: A CA Ac2: C CA S: C II 3	Self care book, usual care as needed
Facco 2007* [39]	Migraine	160 (21%)	Adequate†	Migraine Disability Index (MIDAS) at 3 months	N: 20 D: 30 T: 11w	Ac: A CA S1: II A 3 S2: II C 3	Rizatriptan for attacks in all patients
Foster 2007 [41]	Osteoarthritis of the knee	352 (7%)	Adequate	WOMAC pain scale at 6 weeks (PMOM: 6 months)	N: 6 D: 30 T: 3w	Ac: A CA S: II A 3	Individ. exercise, advice, NSAIDs if needed (I)
Helms 1987 [44]	Primary dysmenorrhea	48 (10%)	Unclear	Monthly pain score week 12	N: 9 D: 30 T: 12w	Ac: C CA S: I C 1	No treatment
Kotani 2001 [46]	Treatment-resistant pain at abdominal scares	70 (0)	Adequate	VAS intensity continous pain, after treatment	N: 20 D: ‡ T: 4w	Ac: A TA S: I A 1	Diclofenac as necessary
Lee 2009 [47]	Chronic prostatitis/ chronic pelvic pain	39 (19%)	Unclear	PMOM: change NIH-Chronic Prostati-tis Symptom Index	N: 12 D: 20 T: 6w	Ac: C EA S: I C 1	Advice and basic exercises
Leibing 2001 [48]	Chronic low back pain	131 (13%)	Unclear	Decrease intensity of pain (VAS) at 12 weeks	N: 20 D: 30 T: 12w	Ac: C CA S: I C 1	26 sessions standardized physiotherapy (I)
Linde 2005 [51]	Migraine	302 (9%)	Adequate	PMOM: Days with at least moderate head-ache weeks in 9 to 12	N: 12 D: 30 T 8w	Ac: B CA Sc: I B 1	Attack medication as needed
Melchart 2005 [53]	Tension-type headache	270 (8)%	Adequate	PMOM: Number of days with headache in weeks 9 to 12	N: 12 D: 30 T: 8w	Ac: B CA S: I B 1	Pain medication as needed
Molsberger 2002 [54]	Chronic low back pain	186 (6%)	Adequate	VAS pain intensity after 1 month (dichotomous PMOM)	N: 12 D: 30 T: 4w	Ac: B CA S: I C 1	Orthopedic rehabilitation program (I)
Suarez-Almazor 2010* [59]	Osteoarthritis of the knee	535 (8%)	Adequate †	WOMAC pain scale at 3 months	N: 12 D: 20 T: 6w	Ac: C CA S: I C 1	NSAIDs and analgesics as before study
Witt 2005 [62]	Osteoarthritis of the knee	294 (5%)	Adequate	PMOM: WOMAC total score after baseline – 8 weeks	N: 12 D: 30 T: 8w	Ac: B CA S: I B 1	NSAIDs if needed

†additional information received from author; *two sham groups with different context; ‡treatment with intradermal needles; §further 25 patients included but not randomized to treatment groups

Column "Outcome measure used for meta-analysis: PMOM: explicitly predefined (regarding *type and timing*) confirmatory main outcome measure, VAS = visual analogue scale; VAS: visual analogue scale

Column "Intervention details": N = number of treatment sessions; D = duration of each treatment session; T: total duration of treatment in weeks; Ac = Acupuncture; S = sham intervention; A = individualized, B = semi-standardized, C = standardized; CA = needling at classical acupuncture body points, EA = electroacupuncture (needles stimulated with electrical currency), EarA = ear acupuncture (needling at ear points), JA = Japanese acupuncture (superficial needling), TA = acupuncture at trigger points; I = penetrating, II = non-penetrating; 1 = needled outside of known points; 2 = acupuncture points not indicated for condition needled; 3 = at correct points

Column "Standard basic care in all groups": (I)= classified as intense co-intervention likely to influence the outcome)

Table 2 - Characteristics of included trials: Short-term trials

Trial	Clinical problem	Sample size (% drop-out	Conceal-ment	Outcome used for meta-analysis	Intervention details		Standard basic care in all groups
Cabrini 2006 [36]	Bronchoscopy	48 (0)	Unclear	VAS discomfort after bronchoscopy	N: 1 D: 20 T: -	Ac: C CA S: I C 1	Lidocaine as needed
Dundee 1986 [38]	Perioperative nausea (minor gyn. Operations)	75 (0)	Unclear	Number of patients vomiting	N: 1 D: 5 T: -	Ac: C CA S: I C 1	Premedication 10 mg nalbophine
Fanti 2003 [40]	Coloscopy	30 (0)	Unclear	Satisfaction with sedations using a verbal rating scale	N: 1 D: 4 T: -	Ac: C EA S: I C 1	Midazolam 0.02 mg/kg before and during colosc. (I)
Gioia 2006 [43]	Sedation during cataract surgery	75 (0)	Unclear	Postoperative anxiety evaluation (VAS)	N: 1 D: 60 T: -	Ac: C CA S: I C 1	Topical eye anesthesia (lidocaine 4%)
Karst 2006 [45]	Anxiety and tooth extraction	48 (0)	Unclear	VAS anxiety	N: 1 D: 25 T: -	Ac: C EarA S: II C 1	Local anesthesia (articaine hydrochloride)
Li 1991 [50]	Coloscopy	36(0)	Unclear	VAS pain intensity	N: 1 D: 30 T: -	Ac: C EA S: I C 1	Routine analgesia and sedation as needed
Rusy 2002 [57]	Postoperative nausea after tonsillectomy	120 (0)	Unclear	Incidence of nausea, vomiting, rescue therapy in first 24 hs	N: 1 D: 20 T: -	Ac: C EA S: I C 2	Standardized anesthesia, analgesia as needed
Ziaei 2006 [64]	Pain reduction and relaxation during labor	90 (unclear)	Unclear	VAS pain after 2 hours	N: 1 D: n.i.. T: -	Ac: C CA S: I C 1	Unclear

See table 1 for legend

Table 3 - Characteristics of included trials: trials on various other conditions

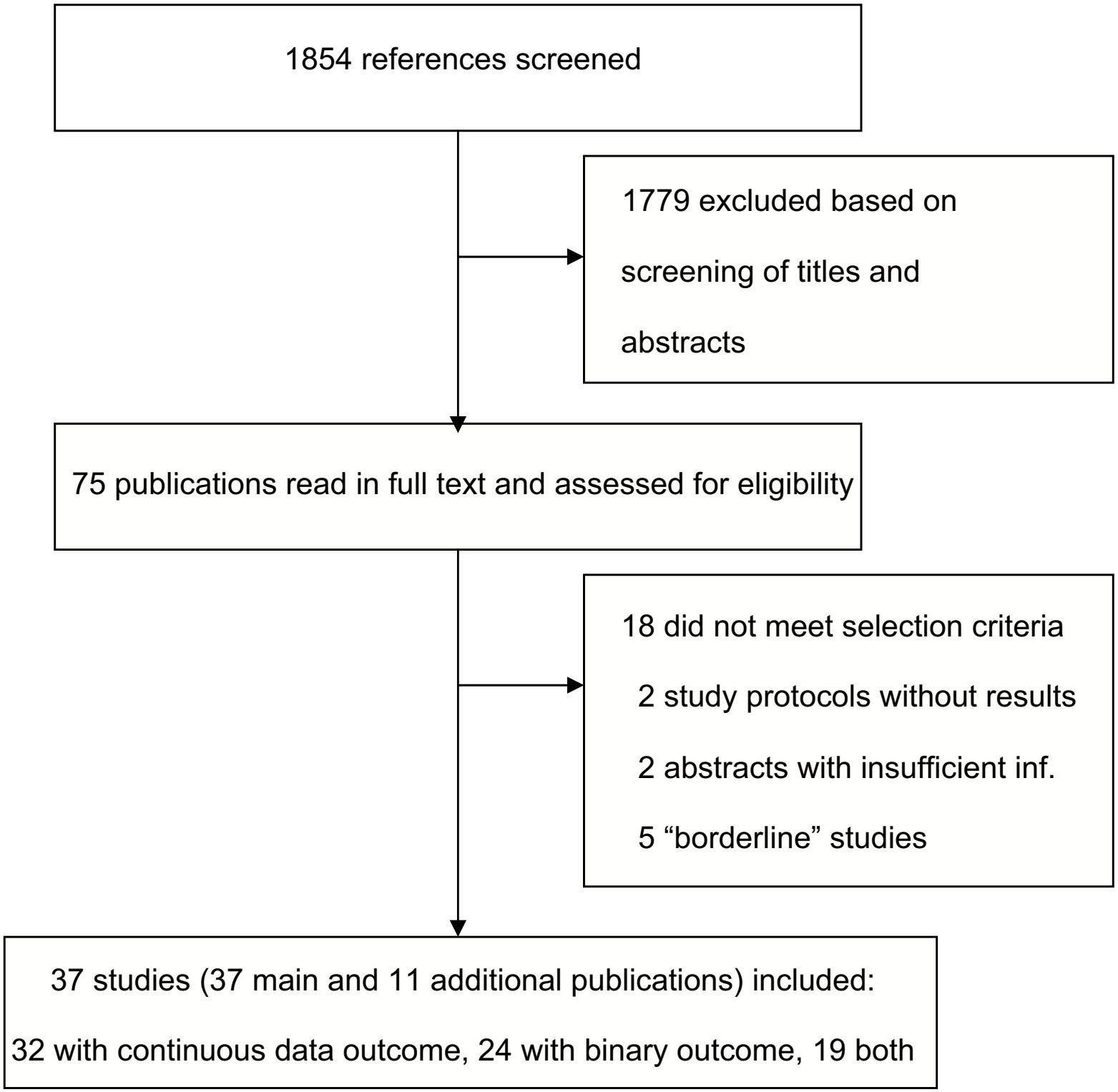
Trial	Clinical problem	Sample size (% drop-out	Conceal-ment	Outcome used for meta-analysis	Intervention details		Standard basic care in all groups
Allen 1998 [28]	Depression	38 (11%)	Unclear	Hamilton Rating Scale for Depression after 8 weeks	N: 12 D: n.i. T: 8w	Ac: A CA S: I A 2	Probably no treatment at all
Allen 2006 [29]	Depression	157 (13%)	Unclear	Hamilton Rating Scale for Depression after 8 weeks (PMOM)	N: 12 D: 20 T: 8w	Ac: A CA S: I A 2	Probably no treatment at all
Asher 2009 [30]	Induction of labor	89 (0)	Adequate	PMOM: Time to delivery	N: 7 D: 30 T: 2w	Ac: C CA S: I A 1	Routine prenatal care
Aune 1998 [31]	Recurrent urinary tract infections	67 (unclear)	Unclear	Patients without infection	N: 8 D: 20 T: 4w	Ac: A CA S: I C 1	No treatment
Avis 2008 [32]	Menopausal hot flashes	56 (0)	Adequate†	Reduction in hot flash frequency†	N: 16 D: 30 T: 8w	Ac: B CA S: B A 2	Continuation of non-drug treatm. used before trial
Bullock 1999 [35]	Addiction – cocaine abuse	236 (59%)	Unclear	Addiction severity scale drug use in week 8	N: 28 D: 45 T: 8w	Ac: C EarA S: I C 2	Eden psychosocial programming (I)
Freire 2006 [42]	Moderate obstructive sleep apnea syndrome	36 (28%)	Unclear	Apnea-hypopnea index after 12 weeks	N: 10 D: 30 T: 10w	Ac: C CA S: I C 1	Offer to receive sleep hygiene counselling
Lembo 2009* [49]	Irritable bowel syndrome (IBS)	231 (8%)	Adequate	PMOM: IBS Global Improvement Scale week 3	N: 6 D: 20 T: 3w	Ac: B CA S: II B 1	Continuation of drugs and diet used before trial
Medici 2002 [52]	Stable chronic asthma	66 (0)	Unclear	PMOM: Peak exp. flow variability baseline – 4 months	N: 16 D: 20 T: 8w	Ac: C CA S: I C 1	Asthma drugs adapted if necessary (I)
Rampes 1997 [55]	Addiction - alcohol abuse	59 (54%)	Adequate †	PMOM: VAS craving for alcohol after 8 weeks	N: 6 D: 30 T: 6w	Ac: C EarA S: I C 2	Individual counselling and group therapy (I)
Röschke 2000 [56]	Depression	70 (unclear)	Unclear	Responder according to Global Assessment Scale (GAS)	N: 12 D: 30 T: 4w	Ac: C CA S: I C 1	Mianserin 90 to 120 mg/day in all groups (I)
Smith 2002 [58]	Nausea & vomiting during pregnancy	445 (25%)	Adequate	Rhodes Index of Nausea after 4 weeks	N: 5 D: 20 T: 4w	Ac: A CA S: I C 1	Pre-trials treatm. continued; life-style recommend.
Tremeau 1992 [60]	Cervical maturation 37th to 38th pregnancy week	98 (18%)	Unclear	PMOM: Bishop score baseline – 10 days	N: 3 D: 20 T: 1w	Ac: C CA S: I C 1	None
Wang 2009 [61]	Low back and pelvic pain during pregnancy	159 (4%)	Adequate†	PMOM: VAS pain change after 1 week†	N: 1 D: ‡ T: 1w	Ac: C EA S: I C 2	Acetaminophen and other self care if needed
Worner 1992 [63]	Addiction - alcohol abuse	56 (95%)	Unclear	Patients with either relapse or in-patient detoxification	N: n.i. D: 30 T: 12w	Ac: C CA S: II C 1	Counselling and group therapy (H)

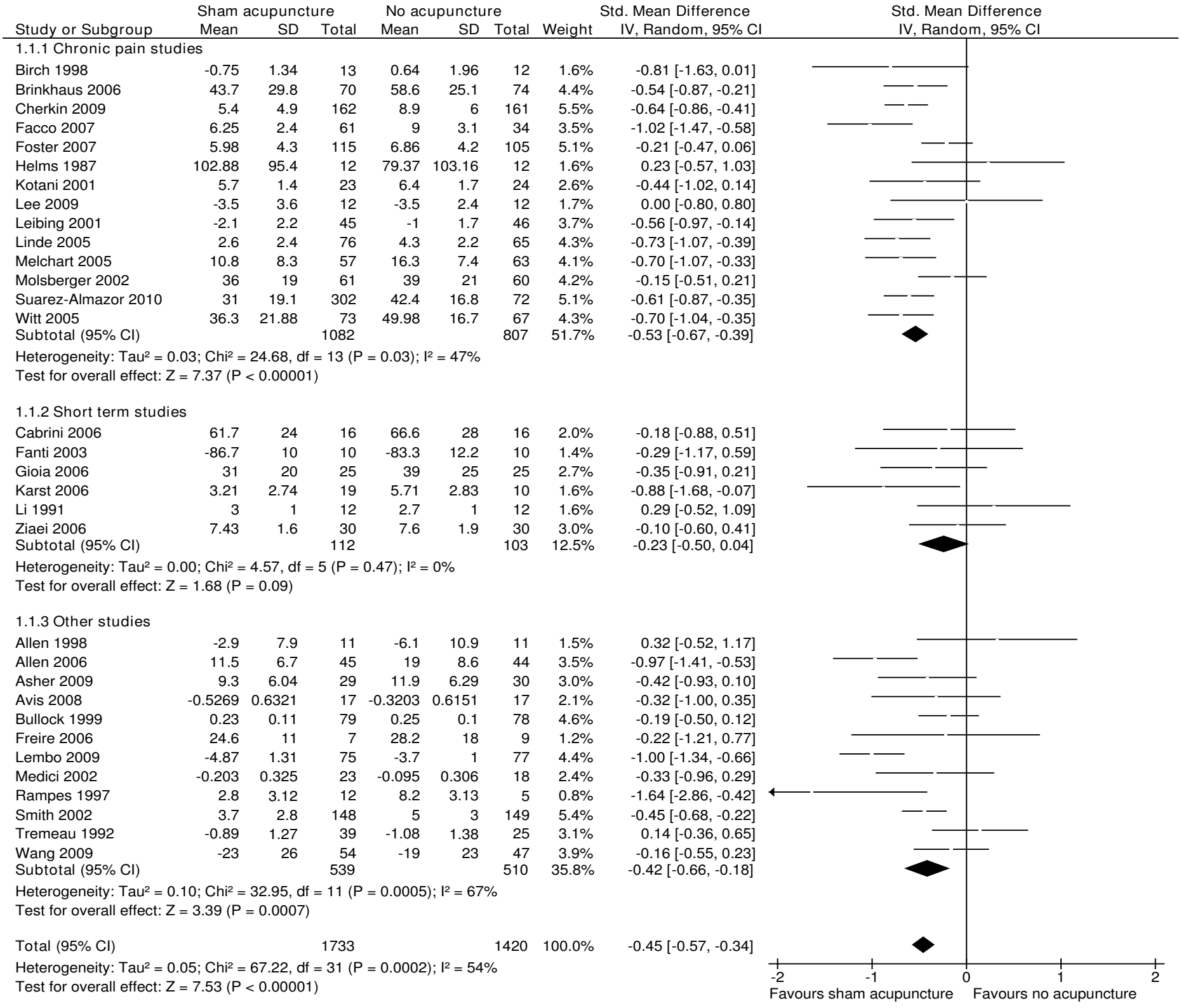
See table 1 for legend

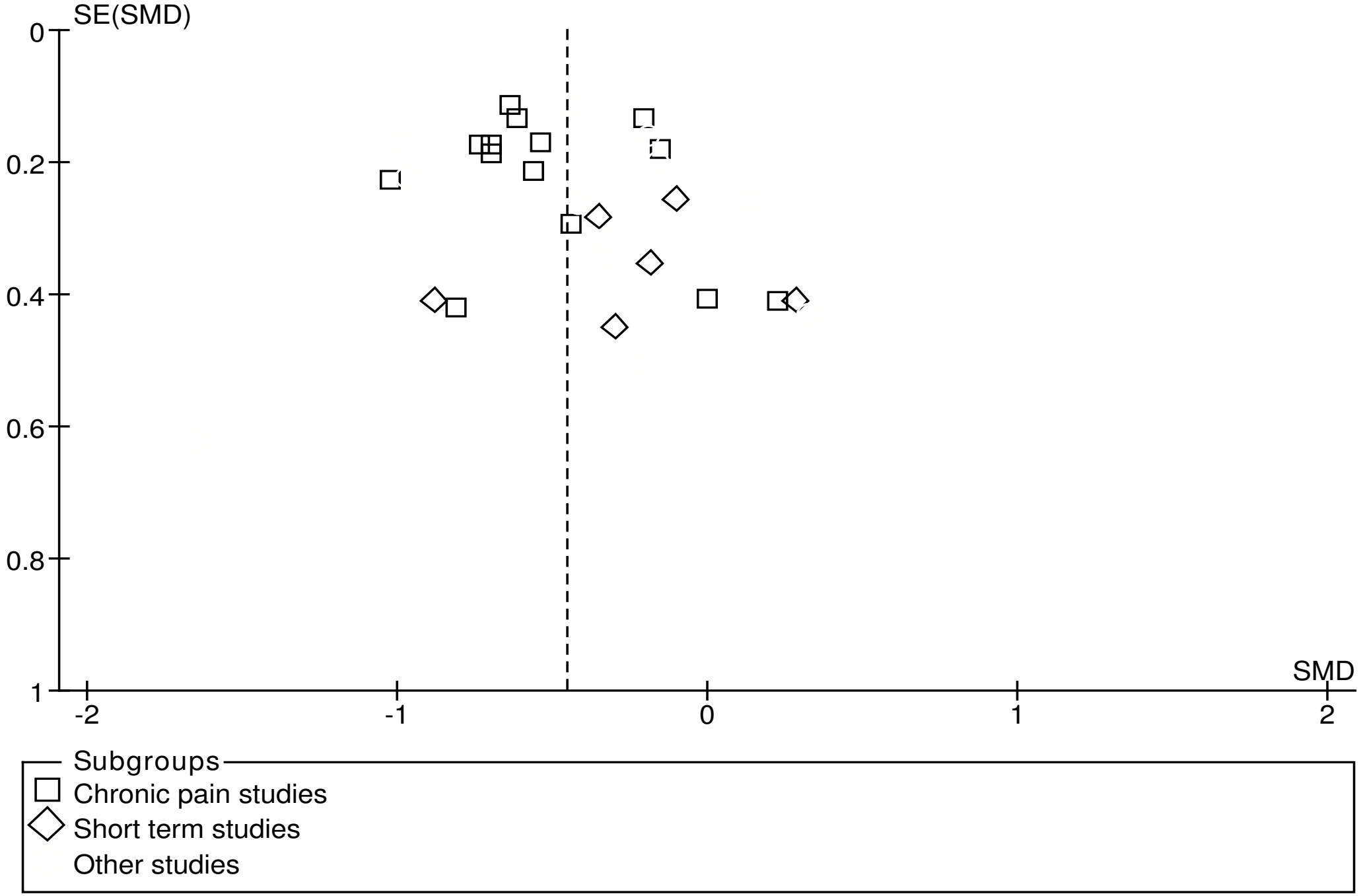
Additional File

Web appendix

Search strategy Medline 19.04.10 (1966-2010 week 15), Search strategy Embase search, 19.04.10 (1988-2010 week 15), Web appendix table 1 - Additional publications related to included studies, Web appendix table 2 - Excluded studies, Web appendix table 3 - Subgroup and sensitivity analyses, Web appendix figure 1 - The “specific” effect of acupuncture (difference between groups receiving acupuncture and sham acupuncture), Web appendix figure 2 - Funnel plot of studies comparing sham acupuncture vs. no acupuncture, Web appendix figure 3 - The “total” effect of acupuncture (difference between groups receiving acupuncture and no acupuncture), Web appendix figure 4 - Funnel plot of studies comparing acupuncture vs. no acupuncture; References appendix


1854 references screened
1779 excluded based on screening of titles and abstracts
75 publications read in full text and assessed for eligibility
18 did not meet selection criteria
2 study protocols without results
2 abstracts with insufficient inf.
5 “borderline” studies
37 studies (37 main and 11 additional publications) included:
32 with continuous data outcome, 24 with binary outcome, 19 both


Sham acupuncture
No acupuncture
Std. Mean Difference
Std. Mean Difference
Study or Subgroup Mean SD Total Mean SD Total Weight IV, Random, 95% CI
IV, Random, 95% CI
1.1.1 Chronic pain studies
Birch 1998 -0.75 1.34 13 0.64 1.96 12 1.6% -0.81 [-1.63, 0.01]
Brinkhaus 2006 43.7 29.8 70 58.6 25.1 74 4.4% -0.54 [-0.87, -0.21]
Cherkin 2009 5.4 4.9 162 8.9 6 161 5.5% -0.64 [-0.86, -0.41]
Facco 2007 6.25 2.4 61 9 3.1 34 3.5% -1.02 [-1.47, -0.58]
Foster 2007 5.98 4.3 115 6.86 4.2 105 5.1% -0.21 [-0.47, 0.06]
Helms 1987 102.88 95.4 12 79.37 103.16 12 1.6% 0.23 [-0.57, 1.03]
Kotani 2001 5.7 1.4 23 6.4 1.7 24 2.6% -0.44 [-1.02, 0.14]
Lee 2009 -3.5 3.6 12 -3.5 2.4 12 1.7% 0.00 [-0.80, 0.80]
Leibing 2001 -2.1 2.2 45 -1 1.7 46 3.7% -0.56 [-0.97, -0.14]
Linde 2005 2.6 2.4 76 4.3 2.2 65 4.3% -0.73 [-1.07, -0.39]
Melchart 2005 10.8 8.3 57 16.3 7.4 63 4.1% -0.70 [-1.07, -0.33]
Molsberger 2002 36 19 61 39 21 60 4.2% -0.15 [-0.51, 0.21]
Suarez-Almazor 2010 31 19.1 302 42.4 16.8 72 5.1% -0.61 [-0.87, -0.35]
Witt 2005 36.3 21.88 73 49.98 16.7 67 4.3% -0.70 [-1.04, -0.35]
Subtotal (95% CI) 1082 807 51.7% -0.53 [-0.67, -0.39]
Heterogeneity: Tau² = 0.03; Chi² = 24.68, df = 13 (P = 0.03); I² = 47%
Test for overall effect: Z = 7.37 (P < 0.00001)
1.1.2 Short term studies
Cabrini 2006 61.7 24 16 66.6 28 16 2.0% -0.18 [-0.88, 0.51]
Fanti 2003 -86.7 10 10 -83.3 12.2 10 1.4% -0.29 [-1.17, 0.59]
Gioia 2006 31 20 25 39 25 25 2.7% -0.35 [-0.91, 0.21]
Karst 2006 3.21 2.74 19 5.71 2.83 10 1.6% -0.88 [-1.68, -0.07]
Li 1991 3 1 12 2.7 1 12 1.6% 0.29 [-0.52, 1.09]
Ziaei 2006 7.43 1.6 30 7.6 1.9 30 3.0% -0.10 [-0.60, 0.41]
Subtotal (95% CI) 112 103 12.5% -0.23 [-0.50, 0.04]
Heterogeneity: Tau² = 0.00; Chi² = 4.57, df = 5 (P = 0.47); I² = 0%
Test for overall effect: Z = 1.68 (P = 0.09)
1.1.3 Other studies
Allen 1998 -2.9 7.9 11 -6.1 10.9 11 1.5% 0.32 [-0.52, 1.17]
Allen 2006 11.5 6.7 45 19 8.6 44 3.5% -0.97 [-1.41, -0.53]
Asher 2009 9.3 6.04 29 11.9 6.29 30 3.0% -0.42 [-0.93, 0.10]
Avis 2008 -0.5269 0.6321 17 -0.3203 0.6151 17 2.1% -0.32 [-1.00, 0.35]
Bullock 1999 0.23 0.11 79 0.25 0.1 78 4.6% -0.19 [-0.50, 0.12]
Freire 2006 24.6 11 7 28.2 18 9 1.2% -0.22 [-1.21, 0.77]
Lembo 2009 -4.87 1.31 75 -3.7 1 77 4.4% -1.00 [-1.34, -0.66]
Medici 2002 -0.203 0.325 23 -0.095 0.306 18 2.4% -0.33 [-0.96, 0.29]
Rampes 1997 2.8 3.12 12 8.2 3.13 5 0.8% -1.64 [-2.86, -0.42]
Smith 2002 3.7 2.8 148 5 3 149 5.4% -0.45 [-0.68, -0.22]
Tremeau 1992 -0.89 1.27 39 -1.08 1.38 25 3.1% 0.14 [-0.36, 0.65]
Wang 2009 -23 26 54 -19 23 47 3.9% -0.16 [-0.55, 0.23]
Subtotal (95% CI) 539 510 35.8% -0.42 [-0.66, -0.18]
Heterogeneity: Tau² = 0.10; Chi² = 32.95, df = 11 (P = 0.0005); I² = 67%
Test for overall effect: Z = 3.39 (P = 0.0007)
Total (95% CI) 1733 1420 100.0% -0.45 [-0.57, -0.34]
Heterogeneity: Tau² = 0.05; Chi² = 67.22, df = 31 (P = 0.0002); I² = 54%
Test for overall effect: Z = 7.53 (P < 0.00001)
-2
-1
0
1
2
Favours sham acupuncture
Favours no acupuncture


SE(SMD)
0
0.2
0.4
0.6
0.8
1
-2
-1
0
1
2
SMD
Subgroups
Chronic pain studies
Short term studies
Other studies

Additional files provided with this submission:

Additional file 1: publ_niemann_bmc_appendix_rev.pdf, 144K
http://www.biomedcentral.com/imedia/1938140233469863/supp1.pdf