Meta-analysis: Acupuncture for osteoarthritis of the knee

Manheimer, Eric; Linde, Klaus; Lao, Lixing; Bouter, Lex M.; Berman, Brian M.

published in
Annals of Internal Medicine
2007

DOI (link to publisher)
10.7326/0003-4819-146-12-200706190-00008

document version
Publisher's PDF, also known as Version of record

Meta-analysis: Acupuncture for Osteoarthritis of the Knee

Eric Manheimer, MS; Klaus Linde, MD, PhD; Lixing Lao, PhD, LAc; Lex M. Bouter, PhD; and Brian M. Berman, MD

Background: Knee osteoarthritis is a major cause of pain and functional limitation.

Purpose: To evaluate the effects of acupuncture for treating knee osteoarthritis.

Data Sources: Cochrane Central Register of Controlled Trials, MEDLINE, and EMBASE databases to January 2007. No language restrictions were applied.

Study Selection: Randomized trials longer than 6 weeks in duration that compared needle acupuncture with a sham, usual care, or waiting list control group for patients with knee osteoarthritis.

Data Extraction: Two authors independently agreed on eligibility, assessed methodological quality and acupuncture adequacy, and extracted outcome data on pain and function measures.

Data Synthesis: Eleven trials met the selection criteria, and 9 reported sufficient data for pooling. Standardized mean differences were calculated by using differences in improvements from baseline between patients assigned to acupuncture and those assigned to control groups. Compared with patients in waiting list control groups, patients who received acupuncture reported clinically relevant short-term improvements in pain (standardized mean difference, −0.96 [95% CI, −1.21 to −0.70]) and function (standardized mean difference, −0.93 [CI, −1.16 to −0.69]). Patients who received acupuncture also reported clinically relevant short- and long-term improvements in pain and function compared with patients in usual care control groups. Compared with a sham control, acupuncture provided clinically irrelevant short-term improvements in pain (standardized mean difference, −0.35 [CI, −0.55 to −0.15]) and function (standardized mean difference, −0.35 [CI, −0.56 to −0.14]) and clinically irrelevant long-term improvements in pain (standardized mean difference, −0.13 [CI, −0.24 to −0.01]) and function (standardized mean difference, −0.14 [CI, −0.26 to −0.03]).

Limitation: Sham-controlled trials had heterogeneous results that were probably due to the variability of acupuncture and sham protocols, patient samples, and settings.

Conclusions: Sham-controlled trials show clinically irrelevant short-term benefits of acupuncture for treating knee osteoarthritis. Waiting list–controlled trials suggest clinically relevant benefits, some of which may be due to placebo or expectation effects.

Ann Intern Med. 2007;146:868-877. www.annals.org
For author affiliations, see end of text.

Osteoarthritis is the leading cause of disability among older adults (1, 2). The joint most commonly affected by osteoarthritis is the knee (3, 4). The prevalence, disability, and associated costs of knee osteoarthritis are expected to steadily increase over the next 25 years because of aging in the population (5).

Nonsteroidal anti-inflammatory drugs (NSAIDs) and acetaminophen are the most commonly used pharmacologic agents for treating knee osteoarthritis (6, 7). However, according to a recent systematic review (8), NSAIDs are only slightly better than placebo in providing short-term pain relief and their effects are probably too small to be meaningful to patients (8). Furthermore, many NSAIDs are associated with considerable side effects (9). Gastrointestinal bleeding, the most clinically substantial effect (10), causes approximately 2200 deaths and 12 000 emergency hospital admissions each year in the United Kingdom alone (11) and is of particular concern to older patients (10). Acetaminophen may have a better toxicity profile than that of NSAIDs (6). However, a recent systematic review (12) suggests that acetaminophen is modestly less effective than NSAIDs and that the clinical significance of acetaminophen is questionable because it results in only a 5% greater improvement from baseline in pain than does placebo in the short term.

The evidence for nonpharmacologic treatments for knee osteoarthritis is generally sparse and inconclusive (13). However, 2 effective nonpharmacologic treatments are exercise (14) and weight loss (15). Some patients with osteoarthritis, however, may have difficulty exercising or losing weight.

The need for additional safe and effective treatments for osteoarthritis is clear. Acupuncture is a safe treatment that has a low risk for serious side effects (16–19). Given its safety, whether acupuncture is effective for treating osteoarthritis of the knee is a highly relevant question.

Our objective was to conduct a systematic review and meta-analysis of the effects of acupuncture for treating knee osteoarthritis.

See also:

Print
Editors' Notes 869

Web-Only
Appendix Tables
Conversion of figures and table into slides

METHODS

Data Sources and Study Selection

We searched the MEDLINE, EMBASE, and Cochrane Central Register of Controlled Trials databases to January 2007 to identify randomized, controlled trials (RCTs). We combined acupuncture-related terms with osteoarthritis-related terms and limited the search to RCTs (20). We

considered older RCTs that were included in previous reviews of acupuncture for osteoarthritis (21–24) for inclusion.

Two authors independently selected articles and resolved disagreements by discussion. Our selection criteria were published RCTs of acupuncture in patients who had received a diagnosis of knee osteoarthritis. We considered the 2 outcomes of pain and function. We applied no language restrictions.

We included only RCTs in which the acupuncture treatment involved the insertion of needles into traditional meridian points. The needles could be inserted into tender points in addition to the traditional meridian points and could be electrically stimulated. We excluded RCTs of dry needling or trigger-point therapy. We also excluded RCTs that compared only 2 different forms of active acupuncture.

We prespecified that trials have at least 6 weeks of observation. This criterion has not been validated as a threshold for study inclusion. However, we thought that RCTs with observation periods less than 6 weeks may also have methodological shortcomings (25, 26) that may exaggerate their results of benefits (27, 28).

Data Extraction and Quality Assessment

Two authors independently extracted data and resolved disagreements by discussion. They extracted information pertaining to the quality of the methods, participants, interventions, and outcomes (including adverse effects). When a study reported more than 1 pain or function outcome measure, we gave preference to the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) pain and function measures because the WOMAC is the most widely used and thoroughly validated instrument for assessing patients with knee osteoarthritis (29–31). We contacted corresponding authors via e-mail and requested that they review the information we extracted from their studies, clarify any ambiguities, and supply missing information.

We evaluated internal validity of the RCTs by using an 11-item scale developed by the Cochrane Back Review Group (32). We considered a score of 6 or more points to indicate high internal validity on the basis of data from ongoing research, in which 1 of the authors is involved. For the patient and outcomes assessor blinding items on the scale, we assigned sham-controlled trials 0.5 point rather than 1 point because we could not be certain that all shams were sufficiently credible in fully blinding patients to the treatment being evaluated. However, we assigned 1 point to sham-controlled trials that evaluated the credibility of the sham and found it to be indistinguishable from true acupuncture.

Acupuncture Assessment

No consensus exists on how best to assess treatment adequacy in acupuncture RCTs, and no methods have been validated (26). We used a method that involved assessing the adequacy of the following 4 aspects of the acupuncture treatment: choice of acupuncture points, number of sessions, needling technique, and experience of the acupuncturist. The adequacy of the sham intervention was also assessed by using an open-ended question. Two acupuncturists, who had previously used this adequacy assessment instrument for an earlier systematic review (33) on acupuncture, made these assessments.

The acupuncturists assessed adequacy independently and achieved consensus by discussion. Assessments were based on only the description of the study population and the acupuncture procedure. The assessors were blinded to the results of the study and the publication. We asked the assessors to guess the identity of each study being assessed to test the success of the blinding.

Data Synthesis and Analysis

We placed RCTs into categories according to control groups, which were sham, usual care, and waiting list. We defined sham control as a sham intervention that was designed to be credible as the active treatment. We defined the usual care control as groups that received some additional standard care therapy that was not provided to the acupuncture group and waiting list control as groups that received no care while waiting for acupuncture.

For our meta-analyses, we defined the short-term follow-up point as the measurement point closest to 8 weeks but no longer than 3 months after randomization. We defined the long-term outcome as the measurement point closest to 6 months but longer than 3 months after randomization.

Because some RCTs used the visual analogue scale version of the WOMAC instrument and others used the Lik-

Context

Previous studies have come to inconsistent conclusions about the effectiveness of acupuncture for treating knee osteoarthritis.

Contribution

This meta-analysis of 9 trials showed that sham-controlled trials identified no clinically meaningful short-term benefits in pain or function with acupuncture for knee osteoarthritis, although trials that did not use a sham control identified some benefits.

Implications

The use of different types of comparisons (sham acupuncture vs. interventions in which the participant knew whether they were receiving acupuncture) explains the variability in the conclusions of published trials about the effectiveness of acupuncture for treating knee osteoarthritis. Placebo or expectation effects probably account for the observed benefits.

—The Editors

Table. **Characteristics of Studies Included in Systematic Review***

Study, Year (Reference)	Location	Mean Age, *y*	Men, %	Internal Validity		Acupuncture			
				11 Items§	Composite Score‖	Style¶	Sessions, *n*	Duration, *wk*	Patients, *n***
Berman et al., 1999 (36)	United States	65	40	1, 1††, 1, 0, 0, 0, 1, 1, 1, 1, 1	8	Formula with EA	16	8	36
Berman et al., 2004 (43)	United States	65	36	1, 1, 1, 0.5, 0, 0.5, 1, 1, 1, 1, 1 1, 1, 1, 0, 0, 0, 1, 0, 0, 1, 1	9 6	Formula with EA	23	26	190
Christensen et al., 1992 (41)	Denmark	69	31	1††, 1††, 1, 0, 0, 1/0‡‡, ?, ?, 1, 1, 0	5/6	Formula acupuncture	6	3	17
Molsberger et al., 1994 (42)	Germany	60	37	?, 0, 1, 0.5, 0, 0.5, 1, ?, 0, 1, 0	4	Formula acupuncture	10	5	71
Sangdee et al., 2002 (44)	Thailand	63	22	?, 0, 1, 0.5, 0, 0.5, 1, 1, 1, 1, 0 ?, 0, 1, 0.5, 0, 0.5, 1, 1, 1, 1, 0	6 6	Formula with EA	12	4	49 48
Scharf et al., 2006 (45)	Germany	63	31	1, 1, 1, 1, 0, 1, 0, 1, 1, 1, 1 1, 1, 1, 0, 0, 0, 0, 1, 1, 1, 1	9 7	Flexible formula	10§§	6	330
Takeda and Wessel, 1994 (35)	Canada	62	50	1††, 1††, 0, 0.5, 0, 0.5, ?, ?, 1, 1, 0	5	Formula	9	3	21
Tukmachi et al., 2004 (37)	United Kingdom	62	17	1, 1, ?, 0, 0, 0, ?, ?, 1, 1, 0 1, 1, ?, 0, 0, 0, ?, ?, 1, 1, 0	4 4	Formula Formula‖ ‖	10 10	5 5	10 10
Vas et al., 2004 (46)	Spain	67	16	1, 1, 1, 0.5, 0, 0.5, 1, 1††, 0, 1, 1	8	Formula with EA	12	12	48
Witt et al., 2005 (47)	Germany	64	34	1, 1, 1, 1, 0, 1, 1, 1, 1, 1, 1 1, 1, 1, 0, 0, 0, 1, 1, 1, 1, 1	10 8	Flexible formula	12	8	150
Witt et al., 2006 (48)	Germany	61	40	1, 1, 1, 0, 0, 0, ?, ?, 1, 1, 1	6	Individualized	11	13	175

* EA = electric stimulation acupuncture; NSAID = nonsteroidal anti-inflammatory drug; RCT = randomized, controlled trial; WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index.

† For all sham-controlled RCTs, the schedule for the sham acupuncture procedure was the same as that of the true acupuncture procedure. For all waiting list–controlled trials, patients on the waiting lists were allowed to receive the current level of oral NSAID or analgesic therapy.

‡ For the 5 RCTs that used a waiting list control group (36, 37, 41, 47, 48), we excluded all outcome measurements after patients on the waiting list began acupuncture. Sangdee and colleagues' trial (44) also had a measurement time point at 9 wk for responders only. All included RCTs reported results as WOMAC scores at all measurement time points listed.

§ Eleven Cochrane Back Review Group criteria: C1 = method of randomization adequate?; C2 = treatment allocation concealed?; C3 = groups similar at baseline regarding the most important prognostic indicators?; C4 = patient blinded to the intervention?; C5 = care provider blinded to the intervention?; C6 = outcome assessor (i.e., the patient for self-assessed WOMAC questionnaire) blinded to the intervention?; C7 = co-interventions avoided or similar?; C8 = compliance acceptable in all groups?; C9 = dropout rate described and acceptable?; C10 = timing of outcome assessment identical for all intervention groups and for all important outcome assessments?; C11 = analysis includes an intention-to-treat analysis. "1" indicates that the criterion was adequately met; "0" indicates that the criterion was not met; "0.5" indicates that the criterion was partially met (for patient and outcomes assessor blinding criteria only); and "?" indicates that the criterion was not reported or was unclear.

‖ Cochrane Back Review Group composite quality score (range, 0–11; score ≥6 indicates good quality).

¶ For the formula style, the same fixed points were used for all patients. For the flexible formula, a fixed formula was used and some additional points were chosen according to the symptoms of the patient.

** For 2 RCTs (41, 42), the number of patients randomly assigned was not clear. For 1 trial (41), we assumed the number of patients randomly assigned on the basis of context. In the other trial (42), we report the number of patients analyzed rather than randomly assigned.

†† Data obtained from authors.

‡‡ Score of patient-assessed outcomes/score of blinded observer-assessed objective outcomes.

§§ For this RCT (45), 5 additional acupuncture sessions were administered if the treatment was graded as partially successful after 6 wk.

‖ ‖ Patients in the second acupuncture group of this RCT (37) were prohibited from taking current NSAID or analgesic medication. We excluded this group from the meta-analysis because patients receiving acupuncture in all other RCTs were not prohibited from taking current medications.

ert version, we used standardized mean differences as the principal measure of effect size so that the results of the RCTs could be combined. We calculated standardized mean differences (Hedge adjusted g) for all RCTs by using differences in improvements between groups divided by the SDs of improvements pooled from the 2 groups (34). For 3 RCTs (35–37), we made some conservative assumptions to compute the standardized mean differences (**Appendix Table 1**, available at www.annals.org).

We used the DerSimonian and Laird (38) model, which is the random-effects model used in RevMan software, version 4.2 (Nordic Cochrane Centre, Copenhagen, Denmark) (39). This model estimates the average treatment effect by incorporating heterogeneity among clinically diverse trials with different, but related, treatment effects (40). When heterogeneity exists, the model (38) assigns smaller studies more weight than they would receive in a fixed-effects model (40). To evaluate statistical heterogeneity within our trial categories, we used I^2 tests on all outcomes included in our meta-analysis.

We conducted sensitivity analyses for the short-term outcome of sham-controlled RCTs, restricting the analyses to RCTs assessed as adequate based on each item of the 11-item Cochrane scale (32) and the 4 aspects of the acupuncture treatment adequacy. We performed additional sensitivity analyses for funding source (industry vs. nonindustry) and follow-up length (≥3 months).

We also evaluated whether the pooled effects of acupuncture met the threshold for minimal clinically important differences, defined as the smallest differences in scores that patients would perceive to be beneficial (29). The clinically relevant effects for knee osteoarthritis have been esti-

Table—Continued

Control†		Co-intervention	Measurement Time Points, wk‡
Type	Patients, n**		
Waiting list	37		4, 8, 12
Sham: combination of penetrating and nonpenetrating needles	191		4, 8, 14, 26
Six 2-h group sessions on arthritis self-management; educational material	189		
Waiting list	15		3, 5, 6, 7
Sham: needles inserted away from acupuncture points	26		5, 13
Sham: patch electrodes attached at knee, with mock EA	49	Diclofenac	4
Sham: patch electrodes attached at knee, with mock EA	47	Placebo diclofenac	
Sham: needles inserted superficially away from acupuncture points	367	Physical therapy	13, 26
10 physician visits with consultation and prescription for diclofenac	342		
Sham: needles inserted superficially 1 inch from acupuncture points	21		3, 7
Waiting list	10		5
Sham: nonpenetrating needles placed at acupuncture points, with mock EA	49	Diclofenac	13
Sham: needles inserted superficially away from acupuncture points	76		8, 26, 52
Waiting list	74		
Waiting list	167		13

mated to be standardized mean differences of 0.39 for WOMAC pain and 0.37 for WOMAC function (29).

Role of the Funding Source

The study was funded by the National Institutes of Health, National Center for Complementary and Alternative Medicine. The funding source had no role in the design, conduct, or reporting of the study or in the decision to submit the manuscript for publication.

RESULTS

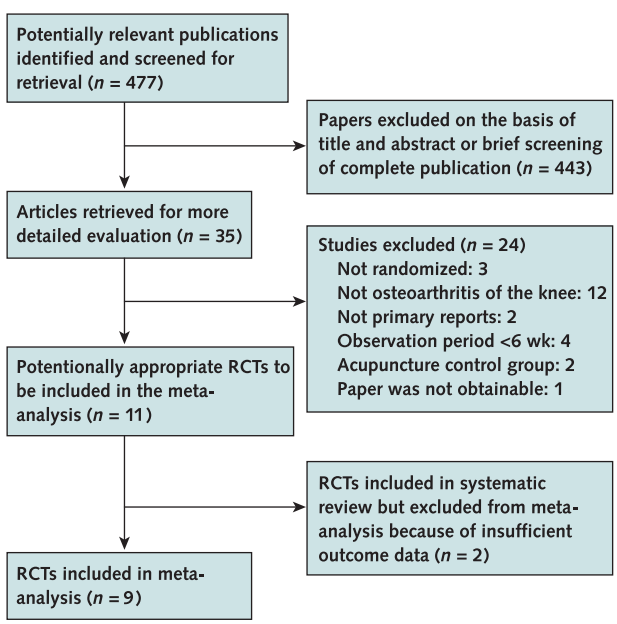
Study Characteristics

We included 11 RCTs (35–37, 41–48) of 2821 patients with osteoarthritis (**Figure 1**). All studies but 1 (42) were published in English. We obtained unpublished data from 8 authors (outcome data [43, 45–48] and methodological information [35, 36, 41, 46]).

The **Table** shows the most important characteristics of the 11 RCTs. All RCTs included patients with a mean duration of osteoarthritis knee pain of 5 years or more. For all RCTs, patients had to have received a diagnosis of knee osteoarthritis to be eligible. All but 1 study (44) required radiologic evidence of the condition. No RCTs reported that the diagnosis was made according to the principles of traditional Chinese medicine.

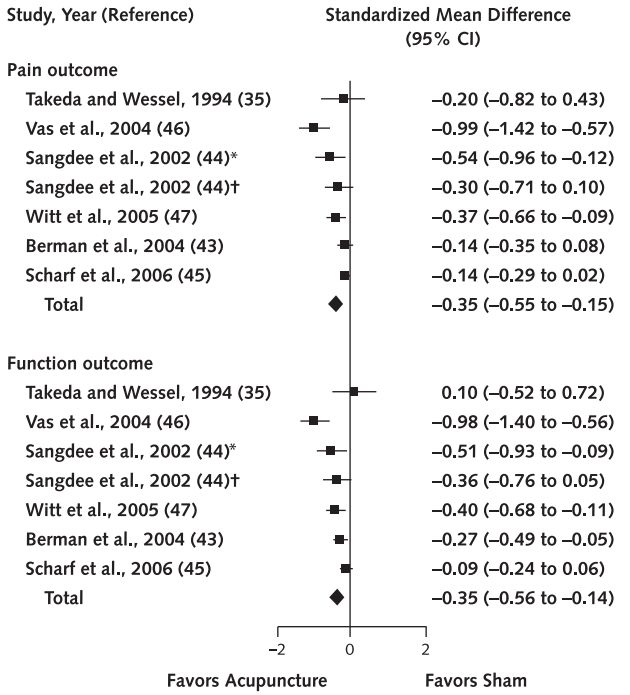
Two RCTs (45, 47) used a flexible formula for point selection, and 8 RCTs (35–37, 41–44, 46) used a set formula. For the 1 remaining pragmatic trial (48), the point selection and needling technique were entirely at the discretion of the treating physicians. Superficial needling alone was used in 1 trial (44), whereas 9 trials (35–37, 41–43, 45–47) used sufficiently deep needle stimulation to elicit the de qi needling sensation. Four trials (36, 43, 44, 46) used electrical stimulation of the needles, 2 of which (44, 46) used electrical stimulation of all local nee-

Figure 1. **Study flow diagram.**



RCT = randomized, controlled trial.

Figure 2. **Effects of acupuncture compared with a sham control group at the short-term measurement point.**



Point estimates and 95% CIs are shown for the standardized mean difference in each study. Summary estimates of the standardized mean differences and their 95% CIs are given by random-effects (DerSimonian and Laird) models. I^2 values were 66% for the pain outcome and 69% for the function outcome. *Comparison of electroacupuncture with sham acupuncture using a diclofenac co-intervention. †Comparison of electroacupuncture with sham acupuncture using a placebo diclofenac co-intervention.

dles. We assessed acupuncture as adequate in the choice of acupuncture points, number of sessions, and needling technique for all trials except for the 1 pragmatic trial (48), in which point selection and needling technique were individualized and therefore could not be assessed (**Appendix Table 2**, available at www.annals.org). The assessors commented that 2 trials (45, 47) used an intensive sham needling technique that may have had physiologic effects. These adequacy assessments were partially unblinded because the assessors were already aware of the identity and results of 3 trials (36, 43, 45).

Among the 7 RCTs that included a sham control, we considered the 4 (43, 45–47) with the highest internal validity ratings on the Cochrane Back Review Group scale (**Table**) to make up the bulk of the evidence for our review. All 4 RCTs were published since 2004 in the world's leading health care journals. Three of the 4 also included a nonacupuncture control group (43, 45, 47). Only 2 of the 4 had any obvious methodological flaws, which were due to higher dropout rates in the sham group for 1 RCT (46) and in the usual care group for the other RCT (43).

Four RCTs did not report funding sources (35, 37, 41, 42). Three (36, 43, 46) were funded by government grants, 1 (44) by a university, and 3 (45, 47, 48) by health insurance companies.

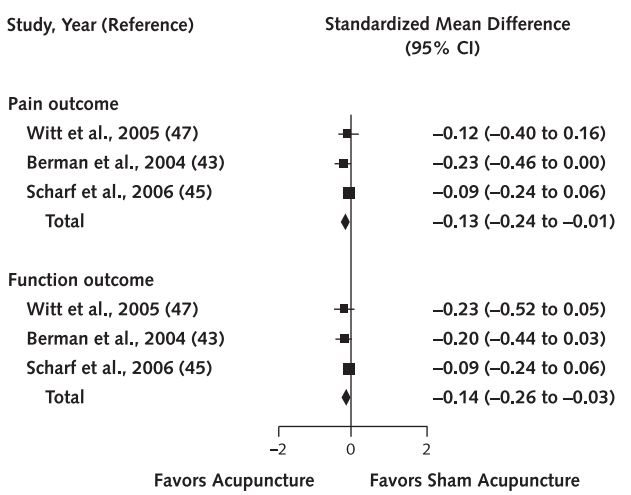
Efficacy of Acupuncture

Compared with a sham control, acupuncture provided clinically irrelevant short-term improvements in pain (standardized mean difference, −0.35 [95% CI, −0.55 to −0.15]) and function (standardized mean difference, −0.35 [CI, −0.56 to −0.14]), but the results were heterogeneous (**Figure 2**). Compared with sham acupuncture, acupuncture also provided clinically irrelevant improvements at 6 months after baseline (**Figure 3**). Patients receiving acupuncture reported clinically relevant short-term improvements compared with patients in waiting list and usual care control groups (**Figure 4**). These clinically relevant improvements were largely maintained at 6 months for trials with usual care control groups (**Figure 5**).

Safety of Acupuncture

Only 3 RCTs (43, 45, 47) described adverse events across groups in detail, and they found that the frequency of adverse events was similar between the acupuncture and control groups (**Appendix Table 3**, available at www.annals.org). Pooling of adverse events across these 3 RCTs was not possible because of limited reporting and heterogeneous methods. No serious adverse events were reported to be associated with acupuncture. The frequency of minor side effects of acupuncture, primarily minor bruising and bleeding at needle insertion sites, ranged from 0% (36) to 45% (44). These frequencies varied widely because of het-

Figure 3. **Effects of acupuncture compared with a sham acupuncture control group at 6 months after baseline.**



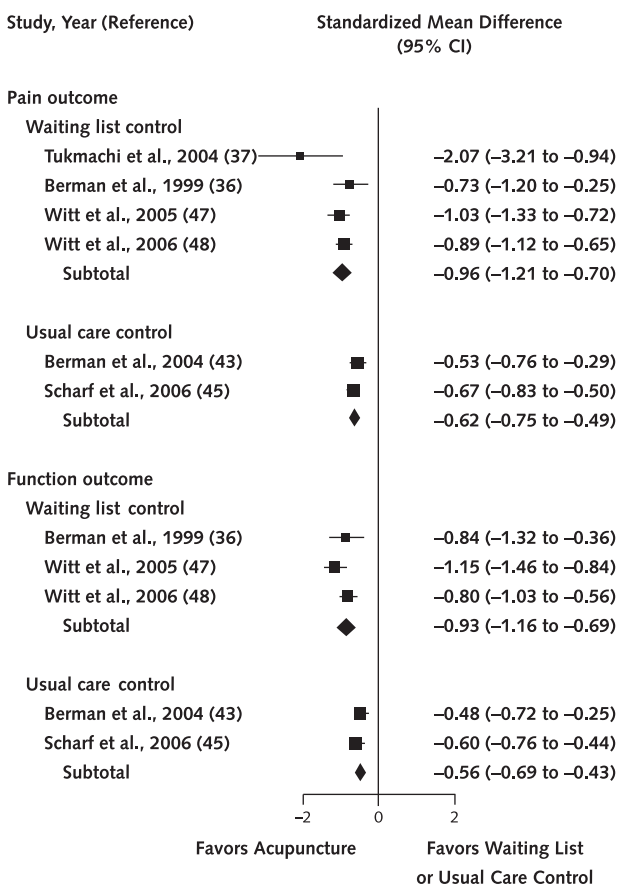
Point estimates and 95% CIs are shown for the standardized mean difference in each study. Summary estimates of the standardized mean differences and their 95% CIs are given by random-effects (DerSimonian and Laird) models. I^2 values were 0% for the pain outcome and 0% for the function outcome.

erogeneous and scanty reporting and different definitions of what constitutes a side effect of acupuncture versus what is an inherent part of treatment (for example, occasional bruising at needle insertion site).

Sensitivity Analysis

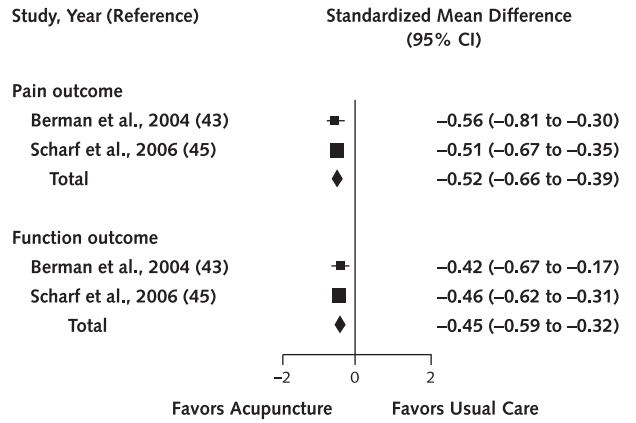
Only 2 of the 17 sensitivity analyses affected the results. These were the analyses on patient blinding and funding source criteria, both of which restricted the analysis to the same 2 RCTs (45, 47) that used superficial needle-penetrating shams. These pooled results were no longer statistically significant for function (standardized mean difference, −0.22 [CI, −0.52 to 0.07]) and were only borderline significant for pain (standardized mean difference, −0.22 [CI, −0.44 to 0.00]).

Figure 4. **Effects of acupuncture compared with a waiting list or usual care control group at the short-term measurement point.**



Point estimates and 95% CIs are shown for the standardized mean difference in each study. Summary estimates of the standardized mean differences and their 95% CIs are given by random-effects (DerSimonian and Laird) models. I^2 values were 42% for waiting list–controlled trials and 0% for usual care–controlled trials for the pain outcome and 39% and 0%, respectively, for the function outcome.

Figure 5. **Effects of acupuncture compared with a usual care control group at 6 months after baseline.**



Point estimates and 95% CIs are shown for the standardized mean difference in each study. Summary estimates of the standardized mean differences and their 95% CIs are given by random-effects (DerSimonian and Laird) models. I^2 values were 0% for the pain outcome and 0% for the function outcome.

DISCUSSION

Current evidence from several large-scale, high-quality RCTs suggests that acupuncture may be an effective treatment for older patients with osteoarthritis of the knee. However, drawing general conclusions is complicated because the effects of acupuncture differ depending on whether acupuncture is compared with a waiting list, usual care, or sham control. The pooled effects of acupuncture were clinically relevant when compared with the waiting list and usual care controls but were not when compared with the sham control.

The clinically relevant effects of acupuncture when compared with the usual care and waiting list controls might be partly attributable to placebo effects of the acupuncture setting (for example, context factors, talking and listening, and credibility of the intervention) (49). Indeed, the fact that both the acupuncture and sham groups reported greater improvements than those of the usual care control groups (43, 45) suggests that acupuncture may elicit a greater placebo effect or meaning response (50) than usual care therapies.

Patient preferences and expectations may also partly explain the findings. In both usual care–controlled trials (43, 45), for example, more patients discontinued treatment in the usual care groups than in either the acupuncture or sham acupuncture group, and most dropouts occurred immediately after treatment assignment. These differences in dropout rates suggest that at least some patients had prerandomization preferences for acupuncture. Any such pretreatment preferences may have influenced later assessments of outcomes (51). Patient expectations that acupuncture will work may also affect patients on a

waiting list. By having to wait for a treatment that they believe is effective (48), patients waiting for acupuncture may be disappointed by the delay, which may influence their ratings of subjective outcomes while waiting. However, despite the limitations of the usual care and waiting list comparator designs, these designs may still best approximate the average likely response to acupuncture in clinical practice, in which treatment effects and placebo factors, expectation effects, and patient preferences may operate (52).

A sham comparator is intended to control for patient expectations and placebo effects and thereby can estimate the effects of acupuncture due solely to the point-specific placement of the acupuncture needles. A sham control should preferably be physiologically inert (53) but also credible enough to patients to control for placebo effects. Ideally, investigators should test the credibility of the sham by asking patients to guess the treatment assignment, rate the credibility of the treatment, or both.

For the 2 sham-controlled RCTs (44, 46) that found clinically relevant benefits of acupuncture compared with sham, the credibility of the shams was not tested. In 1 of these RCTs (44), the sham involved patch electrodes that used mock electrical stimulation and were attached to the same 4 knee points used in the true acupuncture group. Because no needles were used, patients randomly assigned to the sham group probably did not believe that they were receiving traditional acupuncture, although they may have believed that they were receiving a credible treatment (54). In the other RCT that found large, clinically relevant benefits of acupuncture (46), the sham needles did not penetrate the skin. Although some shams with nonpenetrating needles have demonstrated credibility among acupuncture-naive patients (55), the credibility of the sham in the RCT (46) was not tested among the acupuncture-naive patients. Thus, we cannot be certain that this nonpenetrating needle sham was believable for all patients. Although some unblinding may have occurred in these 2 RCTs (44, 46), a possible alternative explanation for the clinically relevant benefits observed is the intensive electrical simulation used at all local knee points, which may produce stronger analgesic effects than manual stimulation of needles (56). After excluding these 2 RCTs (44, 46) in a post hoc sensitivity analysis, we found that the results of the other RCTs that compared acupuncture with sham became almost homogeneous and the pooled results remained statistically significant for pain (standardized mean difference, −0.18 [CI, −0.29 to −0.06]) and function (standardized mean difference, −0.20 [CI, −0.36 to −0.04]).

The third large RCT (43) used an innovative sham acupuncture procedure that involved a combination of penetrating and nonpenetrating needles. In the RCT, most patients believed that they had received true acupuncture at the 4-week credibility test, but patients in the sham group were more likely than those in the acupuncture group to believe that they had received sham treatment at the 26-week test. However, at the 26-week test, patients in the sham group may have been more likely to guess that they were receiving a sham treatment because they were experiencing no benefit (57). In fact, asking patients to guess whether they were assigned to a true or sham treatment is controversial because correctly guessing treatment assignment could be highly correlated with a treatment's effectiveness or lack thereof (57, 58).

The other 2 large RCTs (45, 47) used superficial needle-penetrating shams, which involved an average of 10 (45) to 13 (47) needles inserted superficially at nonacupuncture points. These RCTs found no (45) or minimal (47) clinically relevant differences between true and sham acupuncture, yet they found clinically relevant differences between either acupuncture group (that is, true or sham) and the nonacupuncture comparison groups. Although the needle-penetrating shams used in these 2 RCTs were sufficiently similar to true acupuncture to ensure adequate blinding, they were so similar to true acupuncture that they may have had weak physiologic activity and may have not been true placebo controls (53). Although these sham controls were designed to be inert by involving needles placed superficially and away from acupuncture points, in reality, the avoidance of all 400 estimated acupuncture points in the body (59) may be impossible. In fact, the shams used in these 2 RCTs (45, 47) were judged to be probably physiologically active and inappropriate as controls in another recent systematic review (24). Weak physiologic activity of superficial or sham needle penetration is suggested by several lines of research (53), including RCTs showing larger effects of a superficial needle-penetrating acupuncture than those of a nonpenetrating sham control (44, 60), positron emission tomography research indicating that sham acupuncture can stimulate regions of the brain associated with natural opiate production (61), and animal studies showing that sham needle insertion can have nonspecific analgesic effects through a postulated mechanism of "diffuse noxious inhibitory control" (62). Indeed, superficial needle penetration is a common technique in many authentic traditional Japanese acupuncture styles (63).

In these 2 RCTs that used superficial needle-penetrating shams, the findings of no (45) or minimal (47) clinically relevant differences between acupuncture and sham groups and clinically relevant differences between either acupuncture group (that is, true or sham) and the nonacupuncture comparison groups may also be due to the setting and the resulting patient selection. The 2 RCTs were funded by German insurance agencies to determine whether acupuncture should be reimbursable (64). Participating patients and physicians probably knew that patients would not need to pay for their acupuncture treatments out of pocket in the future if these RCTs had positive results (65). This context may have created a possible financial incentive for patients to overstate benefits of the treatment that they believed to be acupuncture, which in

these well-blinded trials (45, 47) were the true and sham acupuncture treatments, and to understate assessments of nonacupuncture controls. Such biased assessments may have contributed to an overestimation of acupuncture effects compared with the nonacupuncture control and a possible underestimation of any difference in effects between true and sham acupuncture.

Our findings that the pooled effects of acupuncture are statistically significantly superior to those of sham treatments agree with the findings of another recent meta-analysis (24). This is encouraging, considering that the 2 meta-analyses used different definitions of short-term and long-term time points and different data for calculating effect sizes (for example, we obtained unadjusted data directly from RCT investigators, where possible, rather than extracting data from publications). Also, the earlier review (24) did not include the large, very recently published pragmatic trial (48) that we included, but did include 3 small, positive RCTs (66–68) that we excluded because of short duration. In the earlier review (24), sensitivity analyses were conducted by using overall quality scores, and these analyses were largely uninformative. Because using overall quality scores can be problematic (69, 70), we used each individual item related to risk for bias for sensitivity analyses in our review. Finally, in the earlier review (24), the authors grouped together RCTs that compared acupuncture with a waiting list or usual care control group but considered this pooled analysis to be weak because of large statistical heterogeneity. In contrast, we prespecified separate comparisons of acupuncture with waiting list and usual care controls and found fairly low heterogeneity. Indeed, our separate analyses on the effects of acupuncture compared with sham, waiting list, and usual care controls and our use of the minimal clinically important difference measure largely informed our discussion and conclusions.

One potential limitation of our review is small study bias by the trend of an inverse association between trial size and effect size (**Figure 2**). Funnel plots and the Egger regression statistical test also suggested possible small study bias, although the underpowered Egger test had only a nonstatistically significant trend. Although we cannot rule out the possibility that small, negative, sham-controlled RCTs were not identified or were never published, such RCTs would probably not have influenced our findings because any such small, negative studies probably would not have met our 6-week minimum duration criterion. Indeed, we excluded 2 small sham-controlled trials (67, 68) because of short duration, but both were positive. Small study bias may also be caused by an overestimation of treatment effects in smaller trials because of their generally lower internal validity (27). However, in our review, the small study bias was suggested among a set of RCTs that were of high internal validity. A possible alternative explanation for the observed association between trial size and effect size in these RCTs is a less thorough implementation of the acupuncture intervention in the larger trials. For example, the largest trial (45), which showed the smallest differences between acupuncture and sham treatment, was conducted at many practices ($n = 315$) by many physician-acupuncturists ($n = 320$). Monitoring adherence to the predefined acupuncture and sham protocols may have been difficult in this context (45). In contrast, the smaller trials with the larger effects (44, 46) were each conducted at 1 site by 1 acupuncturist.

Future sham-controlled RCTs should ideally use physiologically inactive, yet credible, shams (53, 54). For those that assess longer-term outcomes, current RCT results (43, 47) suggest that maintaining monthly acupuncture treatments in the months before the longer-term assessment may be important.

Our findings of no or minimal clinically relevant effects of acupuncture compared with sham but marked clinically relevant effects of acupuncture compared with a waiting list or usual care control group suggest that placebo effects or the elicitation of the meaning response (50) are largely responsible for the effects of acupuncture. However, acupuncture also seems to have a genuine biological effect, suggested by the small short-term improvements in pain and function compared with sham. Although the improvements provided by acupuncture compared with sham were statistically significant when pooled, we found important heterogeneity of the results, which may be due to differences in sham interventions, acupuncture protocols, settings, and varying proficiencies of the acupuncturists. Because of heterogeneity and small effects, current estimates should be regarded as preliminary. It is too soon to recommend acupuncture as a routine part of care for patients with osteoarthritis. At least 3 large and rigorous sham-controlled trials, 2 still ongoing and 1 recently completed but not yet published, will have a major effect on the existing evidence. Safety is best determined with large prospective surveys of practitioners, and 3 such surveys (17–19) show that serious adverse events after acupuncture are rare. No adverse events were associated with acupuncture in our review, although heterogeneous reporting and relatively small sample sizes limit the usefulness of our data. Because acupuncture may have small short-term effects, some clinicians and patients may consider acupuncture as 1 treatment option in a multidisciplinary approach (13) to treating knee osteoarthritis.

From the University of Maryland School of Medicine, Baltimore, Maryland; Technische Universität München, Munich, Germany; and VU University Medical Center, Amsterdam, the Netherlands.

Note: All study data, including full details of the characteristics of each included RCT and all outcome data extracted from each included RCT, are included in a RevMan file, which will be made available to researchers on request for reproducing these results.

Acknowledgments: The authors thank Nancy Min, PhD, for her support in calculating estimated SDs of change from baseline for publications that did not directly report these but did report other data that

could be used to estimate SDs of change; Byungmook Lim, MD, MPH, PhD, for performing data extraction for Witt and colleagues' 2005 RCT (47); Marcos Hsu, LAc, for assessing the adequacy of the acupuncture administered in the RCTs; and Elizabeth Kimbrough Pradhan, PhD, for providing helpful comments on the manuscript. They also thank Claudia Witt, PhD, for providing unpublished data specific to patients with only knee osteoarthritis from Witt and colleagues' 2006 RCT (48), and Birgitte Christensen, MD; Wenlin Lee, PhD; Jorge Vas, MD; and Steffen Witte, PhD, all coauthors of included RCTs, for confirming and providing data related to their respective RCTs.

Grant Support: Mr. Manheimer and Dr. Berman were funded by a grant from the National Institutes of Health, National Center for Complementary and Alternative Medicine (R24 AT001293). Drs. Lao and Berman were also funded by a grant from the National Center for Complementary and Alternative Medicine (P01 AT002605-01A1).

Potential Financial Conflicts of Interest: *Honoraria:* K. Linde (German Medical Acupuncture Society [Deutsche Ärztegesellschaft für Akupunktur], British Medical Acupuncture Society); *Grants received:* L. Lao (National Institutes of Health), K. Linde (Deutsche Angestellten Krankenkasse, Barmer Ersatzkasse, Kaufmännische Krankenkasse, Hamburg Münchener Krankenkasse, Hanseatische Krankenkasse, Gmünder Ersatzkasse, HZK Krankenkasse für Bau- und Holzberufe, Brühler Ersatzkasse, Krankenkasse Eintracht Heusenstamm, Buchdrucker Krankenkasse), B.M. Berman (National Institutes of Health); *Grants pending:* L. Lao (National Institutes of Health).

Requests for Single Reprints: Eric Manheimer, MS, Center for Integrative Medicine, University of Maryland School of Medicine, 2200 Kernan Drive, Kernan Hospital Mansion, Baltimore, MD 21207.

Current author addresses are available at www.annals.org.

References

1. **Peat G, McCarney R, Croft P.** Knee pain and osteoarthritis in older adults: a review of community burden and current use of primary health care. Ann Rheum Dis. 2001;60:91-7. [PMID: 11156538]

2. Prevalence of self-reported arthritis or chronic joint symptoms among adults—United States, 2001. MMWR Morb Mortal Wkly Rep. 2002;51:948-50. [PMID: 12437034]

3. **Felson DT, Zhang Y.** An update on the epidemiology of knee and hip osteoarthritis with a view to prevention. Arthritis Rheum. 1998;41:1343-55. [PMID: 9704632]

4. **Oliveria SA, Felson DT, Reed JI, Cirillo PA, Walker AM.** Incidence of symptomatic hand, hip, and knee osteoarthritis among patients in a health maintenance organization. Arthritis Rheum. 1995;38:1134-41. [PMID: 7639811]

5. **Lethbridge-Cejku M, Schiller JS, BerLethbridge-Cejku M, Schiller JS, Bernadel L.** Summary health statistics for U.S. adults: National Health Interview Survey, 2002. Vital Health Stat 10. 2004:1-151. [PMID: 15791763]

6. **Wegman A, van der Windt D, van Tulder M, Stalman W, de Vries T.** Nonsteroidal antiinflammatory drugs or acetaminophen for osteoarthritis of the hip or knee? A systematic review of evidence and guidelines. J Rheumatol. 2004; 31:344-54. [PMID: 14760807]

7. **Ausiello JC, Stafford RS.** Trends in medication use for osteoarthritis treatment. J Rheumatol. 2002;29:999-1005. [PMID: 12022364]

8. **Bjordal JM, Ljunggren AE, Klovning A, Slørdal L.** Non-steroidal anti-inflammatory drugs, including cyclo-oxygenase-2 inhibitors, in osteoarthritic knee pain: meta-analysis of randomised placebo controlled trials. BMJ. 2004;329:1317. [PMID: 15561731]

9. **McGettigan P, Henry D.** Cardiovascular risk and inhibition of cyclooxygenase: a systematic review of the observational studies of selective and nonselective inhibitors of cyclooxygenase 2. JAMA. 2006;296:1633-44. [PMID: 16968831]

10. **Hernández-Díaz S, Rodríguez LA.** Association between nonsteroidal anti-inflammatory drugs and upper gastrointestinal tract bleeding/perforation: an overview of epidemiologic studies published in the 1990s. Arch Intern Med. 2000;160:2093-9. [PMID: 10904451]

11. **Blower AL, Brooks A, Fenn GC, Hill A, Pearce MY, Morant S, et al.** Emergency admissions for upper gastrointestinal disease and their relation to NSAID use. Aliment Pharmacol Ther. 1997;11:283-91. [PMID: 9146764]

12. **Towheed TE, Maxwell L, Judd MG, Catton M, Hochberg MC, Wells G.** Acetaminophen for osteoarthritis. Cochrane Database Syst Rev. 2006: CD004257. [PMID: 16437479]

13. **Jordan KM, Arden NK, Doherty M, Bannwarth B, Bijlsma JW, Dieppe P, et al.** Standing Committee for International Clinical Studies Including Therapeutic Trials ESCISIT. EULAR Recommendations 2003: an evidence based approach to the management of knee osteoarthritis: Report of a Task Force of the Standing Committee for International Clinical Studies Including Therapeutic Trials (ESCISIT). Ann Rheum Dis. 2003;62:1145-55. [PMID: 14644851]

14. **Fransen M, McConnell S, Bell M.** Exercise for osteoarthritis of the hip or knee. Cochrane Database Syst Rev. 2003:CD004286. [PMID: 12918008]

15. **Christensen R, Bartels EM, Astrup A, Bliddal H.** Effect of weight reduction in obese patients diagnosed with knee osteoarthritis: a systematic review and meta-analysis. Ann Rheum Dis. 2007;66:433-9. [PMID: 17204567]

16. **Cherkin DC, Sherman KJ, Deyo RA, Shekelle PG.** A review of the evidence for the effectiveness, safety, and cost of acupuncture, massage therapy, and spinal manipulation for back pain. Ann Intern Med. 2003;138:898-906. [PMID: 12779300]

17. **White A, Hayhoe S, Hart A, Ernst E.** Adverse events following acupuncture: prospective survey of 32 000 consultations with doctors and physiotherapists. BMJ. 2001;323:485-6. [PMID: 11532840]

18. **MacPherson H, Thomas K, Walters S, Fitter M.** The York acupuncture safety study: prospective survey of 34 000 treatments by traditional acupuncturists. BMJ. 2001;323:486-7. [PMID: 11532841]

19. **Melchart D, Weidenhammer W, Streng A, Reitmayr S, Hoppe A, Ernst E, et al.** Prospective investigation of adverse effects of acupuncture in 97 733 patients. Arch Intern Med. 2004;164:104-5. [PMID: 14718331]

20. **Glanville JM, Lefebvre C, Miles JN, Camosso-Stefinovic J.** How to identify randomized, controlled trials in MEDLINE: ten years on. J Med Libr Assoc. 2006;94:130-6. [PMID: 16636704]

21. **Ezzo J, Hadhazy V, Birch S, Lao L, Kaplan G, Hochberg M, et al.** Acupuncture for osteoarthritis of the knee: a systematic review. Arthritis Rheum. 2001;44:819-25. [PMID: 11315921]

22. **Ernst E.** Acupuncture as a symptomatic treatment for osteoarthritis. A systematic review. Scand J Rheumatol. 1997;26:444-7. [PMID: 9433405]

23. **Kwon YD, Pittler MH, Ernst E.** Acupuncture for peripheral joint osteoarthritis: a systematic review and meta-analysis. Rheumatology (Oxford). 2006;45: 1331-7. [PMID: 16936326]

24. **White A, Foster NE, Cummings M, Barlas P.** Acupuncture treatment for chronic knee pain: a systematic review. Rheumatology (Oxford). 2007;46:384-90. [PMID: 17215263]

25. **Linde K, Vickers A, Hondras M, ter Riet G, Thormählen J, Berman B, et al.** Systematic reviews of complementary therapies - an annotated bibliography. Part 1: acupuncture. BMC Complement Altern Med. 2001;1:3. [PMID: 11513758]

26. **Birch S.** Clinical research on acupuncture. Part 2. Controlled clinical trials, an overview of their methods. J Altern Complement Med. 2004;10:481-98. [PMID: 15253852]

27. **Jüni P, Altman DG, Egger M.** Systematic reviews in health care: Assessing the quality of controlled clinical trials. BMJ. 2001;323:42-6. [PMID: 11440947]

28. **Schulz KF, Chalmers I, Hayes RJ, Altman DG.** Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. JAMA. 1995;273:408-12. [PMID: 7823387]

29. **Angst F, Aeschlimann A, Michel BA, Stucki G.** Minimal clinically important rehabilitation effects in patients with osteoarthritis of the lower extremities. J Rheumatol. 2002;29:131-8. [PMID: 11824949]

30. **Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW.** Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy for patients with osteoarthritis of the hip or knee. J Rheumatol. 1988;15:1833-40. [PMID: 3068365]

31. **Bellamy N.** The WOMAC Knee and Hip Osteoarthritis Indices: development, validation, globalization and influence on the development of the AUSCAN Hand Osteoarthritis Indices. Clin Exp Rheumatol. 2005;23:S148-53. [PMID: 16273799]

32. **van Tulder M, Furlan A, Bombardier C, Bouter L.** Updated method guidelines for systematic reviews in the cochrane collaboration back review group. Spine. 2003;28:1290-9. [PMID: 12811274]
33. **Furlan AD, van Tulder MW, Cherkin DC, Tsukayama H, Lao L, Koes BW, et al.** Acupuncture and dry-needling for low back pain. Cochrane Database Syst Rev. 2005:CD001351. [PMID: 15674876]
34. **Deeks JJ.** Standard statistical algorithms in Cochrane reviews. RevMan analyses help file: statistical calculations. RevMan Analyses [computer program]. Version 1.0 for Windows. In: Review Manager (RevMan). Version 4.2 for Windows. Copenhagen, Denmark: Nordic Cochrane Centre, Cochrane Collaboration; 2003.
35. **Takeda W, Wessel J.** Acupuncture for the treatment for pain of osteoarthritic knees. Arthritis Care Res. 1994;7:118-22. [PMID: 7727550]
36. **Berman BM, Singh BB, Lao L, Langenberg P, Li H, Hadhazy V, et al.** A randomized trial of acupuncture as an adjunctive therapy for osteoarthritis of the knee. Rheumatology (Oxford). 1999;38:346-54. [PMID: 10378713]
37. **Tukmachi E, Jubb R, Dempsey E, Jones P.** The effect of acupuncture on the symptoms of knee osteoarthritis—an open randomised controlled study. Acupunct Med. 2004;22:14-22. [PMID: 15077933]
38. **DerSimonian R, Laird N.** Meta-analysis in clinical trials. Control Clin Trials. 1986;7:177-88. [PMID: 3802833]
39. Review Manager (RevMan) [computer program]. Version 4.2 for Windows. Copenhagen, Denmark: Nordic Cochrane Centre, Cochrane Collaboration; 2003.
40. **Deeks JJ, Higgins JPT, Altman DG, eds.** Analysing and presenting results. In: Higgins JPT, Green S, eds. Cochrane Handbook for Systematic Reviews of Interventions 4.2.4. Section 8.7.4. The Cochrane Library, Issue 3, 2005. Chichester, UK: J Wiley; 2005. Accessed at www.cochrane.org/resources/handbook/hbook.htm on 3 June 2006.
41. **Christensen BV, Iuhl IU, Vilbek H, Bülow HH, Dreijer NC, Rasmussen HF.** Acupuncture treatment for severe knee osteoarthrosis. A long-term study. Acta Anaesthesiol Scand. 1992;36:519-25. [PMID: 1514335]
42. **Molsberger A, Böwing G, Jensen KU, Lorek M.** Acupuncture treatment for the relief of gonarthritis pain: a controlled clinical trial. Der Schmerz. 1994;8:37-42.
43. **Berman BM, Lao L, Langenberg P, Lee WL, Gilpin AM, Hochberg MC.** Effectiveness of acupuncture as adjunctive therapy for osteoarthritis of the knee: a randomized, controlled trial. Ann Intern Med. 2004;141:901-10. [PMID: 15611487]
44. **Sangdee C, Teekachunhatean S, Sananpanich K, Sugandhavesa N, Chiewchantanakit S, Pojchamarnwiputh S, et al.** Electroacupuncture versus diclofenac in symptomatic treatment for osteoarthritis of the knee: a randomized, controlled trial. BMC Complement Altern Med. 2002;2:3. [PMID: 11914160]
45. **Scharf HP, Mansmann U, Streitberger K, Witte S, Krämer J, Maier C, et al.** Acupuncture and knee osteoarthritis: a three-armed randomized trial. Ann Intern Med. 2006;145:12-20. [PMID: 16818924]
46. **Vas J, Méndez C, Perea-Milla E, Vega E, Panadero MD, León JM, et al.** Acupuncture as a complementary therapy to the pharmacologic treatment for osteoarthritis of the knee: randomised controlled trial. BMJ. 2004;329:1216. [PMID: 15494348]
47. **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;366:136-43. [PMID: 16005336]
48. **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;54:3485-93. [PMID: 17075849]
49. **Paterson C, Dieppe P.** Characteristic and incidental (placebo) effects in complex interventions such as acupuncture. BMJ. 2005;330:1202-5. [PMID: 15905259]
50. **Moerman DE, Jonas WB.** Deconstructing the placebo effect and finding the meaning response. Ann Intern Med. 2002;136:471-6. [PMID: 11900500]
51. **Kalauokalani D, Cherkin DC, Sherman KJ, Koepsell TD, Deyo RA.** Lessons from a trial of acupuncture and massage for low back pain: patient expectations and treatment effects. Spine. 2001;26:1418-24. [PMID: 11458142]
52. **Roland M, Torgerson DJ.** What are pragmatic trials? BMJ. 1998;316:285. [PMID: 9472515]
53. **Birch S.** A review and analysis of placebo treatments, placebo effects, and placebo controls in trials of medical procedures when sham is not inert. J Altern Complement Med. 2006;12:303-10. [PMID: 16646730]
54. **White P, Lewith G, Prescott P, Conway J.** Acupuncture versus placebo for the treatment for chronic mechanical neck pain: a randomized, controlled trial. Ann Intern Med. 2004;141:911-9. [PMID: 15611488]
55. **Streitberger K, Kleinhenz J.** Introducing a placebo needle into acupuncture research. Lancet. 1998;352:364-5. [PMID: 9717924]
56. **Ulett GA, Han S, Han JS.** Electroacupuncture: mechanisms and clinical application. Biol Psychiatry. 1998;44:129-38. [PMID: 9646895]
57. **Cherkin DC, Sherman KJ.** Acupuncture and knee osteoarthritis [Letter]. Ann Intern Med. 2005;142:872; author reply 872-3. [PMID: 15897543]
58. **Schulz KF, Grimes DA.** Blinding in randomised trials: hiding who got what. Lancet. 2002;359:696-700. [PMID: 11879884]
59. **World Health Organization**. A proposed standard international acupuncture nomenclature: report of a WHO scientific group. Geneva: World Health Organization; 1991.
60. **Macdonald AJ, Macrae KD, Master BR, Rubin AP.** Superficial acupuncture in the relief of chronic low back pain. Ann R Coll Surg Engl. 1983;65:44-6. [PMID: 6218776]
61. **Pariente J, White P, Frackowiak RS, Lewith G.** Expectancy and belief modulate the neuronal substrates of pain treated by acupuncture. Neuroimage. 2005;25:1161-7. [PMID: 15850733]
62. **Le Bars D, Dickenson AH, Besson JM.** Diffuse noxious inhibitory controls (DNIC). I. Effects on dorsal horn convergent neurones in the rat. Pain. 1979;6:283-304. [PMID: 460935]
63. **Birch S, Ida J.** Japanese Acupuncture: A Clinical Guide. Brookline, MA: Paradigm Publications; 1998.
64. **Linde K.** Should acupuncture be reimbursed? Three acupuncture programmes of German Statutory Health Insurance Funds (SHI) and their scientific evaluation. Z Arztl Fortbild Qualitatssich. 2004;98:467-8. [PMID: 15527186]
65. **Ernst E.** Testing acupuncture for osteoarthritis: pragmatic trials or efficacy studies? Rheumatology (Oxford). 2006;45:125.
66. **Ng MM, Leung MC, Poon DM.** The effects of electro-acupuncture and transcutaneous electrical nerve stimulation on patients with painful osteoarthritic knees: a randomized, controlled trial with follow-up evaluation. J Altern Complement Med. 2003;9:641-9. [PMID: 14629842]
67. **Petrou P, Winkler V, Genti G, Balint G.** Double-blind trial to evaluate the effect of acupuncture treatment on knee osteoarthrosis. Scandinavian Journal of Acupuncture. 1988;3:113-6.
68. **Yurtkuran M, Kocagil T.** TENS, electroacupuncture and ice massage: comparison of treatment for osteoarthritis of the knee. Am J Acupunct. 1999;27:133-40. [PMID: 10729968]
69. **Jüni P, Witschi A, Bloch R, Egger M.** The hazards of scoring the quality of clinical trials for meta-analysis. JAMA. 1999;282:1054-60. [PMID: 10493204]
70. **Greenland S.** Invited commentary: a critical look at some popular meta-analytic methods. Am J Epidemiol. 1994;140:290-6. [PMID: 8030632]

Current Author Addresses: Mr. Manheimer and Drs. Lao and Berman: Center for Integrative Medicine, University of Maryland School of Medicine, 2200 Kernan Drive, Kernan Hospital Mansion, Baltimore, MD 21207.

Dr. Linde: Centre for Complementary Medicine Research, Department of Internal Medicine II, Technische Universität München, Kaiserstrasse 9, 80801 Munich, Germany.

Dr. Bouter: VU University, Executive Board, VU-Windesheim, De Boelelaan 1105, Room 2d-18, 1081 HV Amsterdam, the Netherlands.

71. **Streitberger K, Witte S, Mansmann U, Knauer C, Krämer J, Scharf HP, et al.** Efficacy and safety of acupuncture for chronic pain caused by gonarthrosis: a study protocol of an ongoing multi-centre randomised controlled clinical trial [ISRCTN27450856]. BMC Complement Altern Med. 2004;4:6. [PMID: 15040805]

Appendix Table 1. **Original Data from Included Studies and Assumptions Used to Derive the Meta-analysis Study Data***

Study, Year (Reference)	Measurement Time Point from Baseline, *wk*	Group	Patients, *n*	WOMAC Subscale	Mean Baseline Value (SD)	Mean Posttreatment Value (SD)	*P* Value
Berman et al., 1999 (36)	8	Acupuncture	36	Function	34.56 (12.20)	20.31 (13.26)	0.001
	8	Waiting list	37	Function	36.19 (9.22)	36.14 (10.55)	0.70
	8	Acupuncture	36	Pain	9.58 (3.26)	5.34 (3.62)	0.001
	8	Waiting list	37	Pain	9.78 (2.83)	9.46 (3.56)	0.48
Tukmachi et al., 2004 (37)	5	Acupuncture	10	Pain	12.20 (1.70)	4.40 (4.30)	0.001
	5	Waiting list	10	Pain	12.60 (3.10)	12.70 (3.70)	0.050
Takeda and Wessel, 1994 (35)	7	Acupuncture	20	Pain	19.44 (13.53)	14.01 (12.29)	NR
	7	Sham	20	Pain	21.93 (8.71)	19.44 (18.91)	NR
	7	Acupuncture	20	Function	61.44 (43.15)	48.03 (43.58)	NR
	7	Sham	20	Function	77.80 (36.55)	60.02 (45.85)	NR

* For the 3 RCTs, SDs of change were not completely reported, and some conservative assumptions were made to calculate standardized mean differences for the analysis of changes from baseline. For Berman and colleagues' (36) and Tukmachi and colleagues' (37) trials, we calculated estimated SDs of changes from baseline from their reported within-group change *P* values; we used conservative assumptions when exact *P* values were not reported. For these calculations, we used the reported *P* values and the mean changes in each group to calculate the SEs of change for each group, which were then converted to SDs of change for each group. Takeda and Wessel's study (35) reported no SDs of changes or any statistics (e.g., *P* values) that would allow us to directly calculate them for any outcomes. For that study (35), we used the pre- and posttreatment means and SDs for each group and assumed a conservative within-patient, pretest–posttest correlation of 0.5 to calculate the SDs of change for each group. NR = not reported; WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index.

Appendix Table 2. **Assessments of Adequacy of Acupuncture and Control Protocols***

Study, Year (Reference)	Experience of Acupuncturists	Comments on Control Group	Guess of Study
Berman et al., 1999 (36)	Adequate†	Patients were not blinded; waiting list control was used; no sham or placebo acupuncture was used	Berman et al., 1999 (36)
Berman et al., 2004 (43)	Adequate	Sham or placebo (insertion sham and noninsertion sham) combination method was adequately used; credibility of the blinding to treatment was assessed	Berman et al., 2004 (43)
Christensen et al., 1992 (41)	Unknown (not reported)	Patients were not blinded; waiting list control was used; no sham or placebo acupuncture was used	Uncertain
Sangdee et al., 2002 (44)	Adequate	Transcutaneous electrical nerve stimulation (patch electrodes) was used, which may not be comparable to acupuncture needle intervention; it is hard to tell from the study's Figure B whether needles were involved; no information was given to validate the success of the blinding	Uncertain
Scharf et al., 2006 (45)‡	Adequate	Too many sham acupuncture points were used (n = 10) with minimal insertion, which may produce therapeutic effects; with 315 practices, it is not clear how the principal investigator could ensure that these minimal sham treatments were properly implemented	1 assessor correctly guessed the study
Takeda and Wessel, 1994 (35)	Unknown (years of practice were not provided)	Sham acupuncture (minimal insertion sham) was used; the sham points selected were too close to the real point (1 inch away), and this may have produced a therapeutic effect; the investigators told patients in both groups that they were receiving 2 different types of acupuncture treatments, and therefore patients in both groups thought they received real treatment	Uncertain
Tukmachi et al., 2004 (37)	Unknown (no mention of acupuncturist's training or certifications)	Patients were not blinded; waiting list control was used; no sham or placebo acupuncture was used	Uncertain
Vas et al., 2004 (46)	Unknown§	Placebo acupuncture (noninsertion sham) was used; no information was given to validate the success of the blinding	Uncertain
Witt et al., 2005 (47)	Adequate	Minimal sham acupuncture at nonacupuncture points was used, and as the authors pointed out, this type of intervention may have a physiologic effect; credibility of the blinding to treatment was assessed	Uncertain
Witt et al., 2006 (48)	Unknown‖	Patients were not blinded; waiting list control was used; no sham or placebo acupuncture was used	Uncertain

* Acupuncturists assessed acupuncture as adequate in terms of the choice of acupuncture points, number of sessions, and needling technique for all trials except for the 1 pragmatic trial (48), for which point selection and needling technique were individualized and therefore could not be reported or assessed. The acupuncturists could not assess the adequacy of acupuncture for the German-language study (42).

† The paper did not mention the qualification of the acupuncturist, although the reviewer knows that the acupuncturist was qualified.

‡ Because this trial was not yet published at the time of the adequacy assessment, the assessments were based on the descriptions in the published protocol (Streitberger et al., 2004 [71]). At the time of the assessments, 1 acupuncturist was unaware of the study or its results. The other acupuncturist had heard that the trial found no statistically significant differences between the acupuncture and sham groups.

§ The acupuncturist was a graduate of Beijing University of Medical Sciences, which is a Western medical school. We are not sure whether he or she had adequate acupuncture training. Also, no information on the years of experience was provided.

‖ For the individualized treatment used in this trial, the acupuncturists should ideally have extensive training to be able to make the best selection of acupuncture points and to provide the best needle technique. Although this trial's requirement for ≥140 h of acupuncturist training is perhaps adequate for trials in which a fixed, standardized treatment protocol is used, it may not be adequate in this context, in which the acupuncturist is required to select acupuncture points. However, because 140 h was only a minimal training requirement and because many of the physicians may have had years of practice (years of experience were not reported in the publication), we scored this trial as unknown.

Appendix Table 3. **Adverse Events in the Acupuncture and Control Groups and Minor Side Effects of Acupuncture***

Study, Year (Reference)	Assessment Methods for Adverse Events	Serious Adverse Events per Group, *n/n*†		Assessment Methods for Side Effects of Acupuncture	Side Effects of Acupuncture‡
		Acupuncture	Control		
Berman et al., 1999 (36)	NR	NR	NR	NR	No patients reported side effects from the 16 acupuncture sessions
Berman et al., 2004 (43)	At each measurement time, investigators assessed patients' self-reports of adverse events that were potentially related to acupuncture; assessment instrument and definition of serious adverse event NR	14/190	Sham, 5/191; usual care, 7/189	Participants were asked to report subjective symptoms that could be attributed to acupuncture (such as dizziness, nausea, and numbness) during the study	Changes in subjective symptoms after baseline assessment did not statistically significantly differ among the 3 groups, and the incidence of these symptoms was low throughout the RCT (numbers NR)
Christensen et al., 1992 (41)	NR	NR	NR	NR	Approximately 16% of patients reported minor side effects of acupuncture (e.g., nausea, dizziness, and bruising); no patients withdrew from study because of side effects
Molsberger et al., 1994 (42)	NR	NR	NR	NR	NR
Sangdee et al., 2002 (44)	Nondirective questioning for adverse events performed weekly for 4 wk	0§	0§	Physical examination of the knee for contusions	Approximately 45% of patients in the electroacupuncture groups had local bruises around the knee; however, the bruises usually disappeared within 5–7 d
Scharf et al., 2006 (45)	At each visit, the investigator documented all serious and nonserious events since the last visit; events were coded according to the *Medical Dictionary for Regulatory Activities*	20/330‖	Sham, 9/367; usual care, 16/342‖	Same assessment method as that used for adverse events (i.e., adverse events and minor side effects were not differentiated); all events coded for acupuncture and control groups	Of all events coded, only bruising was reported more often in the acupuncture and sham groups than in the usual care group; approximately 7% of patients in the acupuncture group and 10% of patients in the sham group reported bruising
Takeda and Wessel, 1994 (35)	NR	NR	NR	NR	NR
Tukmachi et al., 2004 (37)	NR	NR	NR	NR	NR
Vas et al., 2004 (46)	NR	NR	NR	NR	3 patients reported bruising at 1 acupuncture point (SP6); it is assumed, but was not explicitly stated, that these 3 patients were among the 48 patients randomly assigned to the true acupuncture group
Witt et al., 2005 (47)	Documented by physicians at each session; assessment instrument and definition of serious adverse event NR; assessments were made at each measurement point, but publication reports only adverse events measured at 26 wk	3/150	Sham, 2/76; waiting list, 4/74	Documented by physicians at each session and reported by patients at the end of wk 8; assessment instrument NR	Approximately 14% of patients in the acupuncture group and 18% of patients in the sham group reported minor side effects, primarily bruising, bleeding, or pain at needle insertion site
Witt et al., 2006 (48)	NR	NR	NR	Side effects were recorded on patient and physician questionnaires after 3 mo	Approximately 5% of patients reported side effects of acupuncture, which were primarily local bleeding, bruising, or pain at needle insertion site; no effects were life-threatening

* Definitions of adverse events and side effects of acupuncture and methods of assessments varied among trials. NR = not reported; RCT = randomized, controlled trial; SP6 = spleen meridian point 6.

† In 2 (43, 47) of the 3 RCT publications that reported number of adverse events per group, it was not clear whether the numbers reported in the publications (included in the table) were the numbers of adverse events in each group or the numbers of patients in each group who had adverse events. For the third RCT (45), numbers reported in this table are the numbers of patients in each group who had adverse events. Of the serious adverse events observed in the acupuncture groups (43, 45, 47), none was interpreted as treatment-related.

‡ The severity or duration of the minor side effects associated with the needling was not systematically reported in any RCT.

§ Percentages of patients who experienced less than serious adverse events (e.g., hypertension, rash, or gastrointestinal symptoms) did not differ between groups (numbers not reported by authors).

‖ The number of patients with serious adverse events. The investigators did not define what was considered a serious adverse event. They also reported "notable" serious adverse events (also undefined), which occurred for 3 patients in the acupuncture group, 1 in the sham group, and 3 in the usual care group.