

A Randomized Trial Comparing Acupuncture, Simulated Acupuncture, and Usual Care for Chronic Low Back Pain

Daniel C. Cherkin, PhD; Karen J. Sherman, PhD; Andrew L. Avins, MD, MPH; Janet H. Erro, RN, MN; Laura Ichikawa, MS; William E. Barlow, PhD; Kristin Delaney, MPH; Rene Hawkes, BA; Luisa Hamilton, MD; Alice Pressman, MS; Partap S. Khalsa, DC, PhD; Richard A. Deyo, MD, MPH

Background: Acupuncture is a popular complementary and alternative treatment for chronic back pain. Recent European trials suggest similar short-term benefits from real and sham acupuncture needling. This trial addresses the importance of needle placement and skin penetration in eliciting acupuncture effects for patients with chronic low back pain.

Methods: A total of 638 adults with chronic mechanical low back pain were randomized to individualized acupuncture, standardized acupuncture, simulated acupuncture, or usual care. Ten treatments were provided over 7 weeks by experienced acupuncturists. The primary outcomes were back-related dysfunction (Roland-Morris Disability Questionnaire score; range, 0-23) and symptom bothersomeness (0-10 scale). Outcomes were assessed at baseline and after 8, 26, and 52 weeks.

Results: At 8 weeks, mean dysfunction scores for the individualized, standardized, and simulated acupuncture groups improved by 4.4, 4.5, and 4.4 points, respectively, compared with 2.1 points for those receiving usual care ($P < .001$). Participants receiving real or simulated acupuncture were more likely than those receiving usual

care to experience clinically meaningful improvements on the dysfunction scale (60% vs 39%; $P < .001$). Symptoms improved by 1.6 to 1.9 points in the treatment groups compared with 0.7 points in the usual care group ($P < .001$). After 1 year, participants in the treatment groups were more likely than those receiving usual care to experience clinically meaningful improvements in dysfunction (59% to 65% vs 50%, respectively; $P = .02$) but not in symptoms ($P > .05$).

Conclusions: Although acupuncture was found effective for chronic low back pain, tailoring needling sites to each patient and penetration of the skin appear to be unimportant in eliciting therapeutic benefits. These findings raise questions about acupuncture's purported mechanisms of action. It remains unclear whether acupuncture or our simulated method of acupuncture provide physiologically important stimulation or represent placebo or nonspecific effects.

Trial Registration: clinicaltrials.gov Identifier: NCT00065585

Arch Intern Med. 2009;169(9):858-866

Author Affiliations: Center for Health Studies, Seattle, Washington (Drs Cherkin, Sherman, and Barlow and Mss Erro, Ichikawa, Delaney, and Hawkes); Division of Research, Northern California Kaiser Permanente, Oakland (Drs Avins and Hamilton and Ms Pressman); Cancer Research and Biostatistics, Seattle (Dr Barlow); Division of Extramural Research and Training, National Center for Complementary and Alternative Medicine, National Institutes of Health, Bethesda, Maryland (Dr Khalsa); and Department of Family Medicine, Oregon Health and Science University, Portland (Dr Deyo).

AMERICANS SPEND AT LEAST \$37 billion annually on medical care for back pain,^{1,2} and our economy suffers another \$19.8 billion in lost worker productivity.³ There is no evidence that escalating expenses for spine care have improved self-assessed health status.²

Many patients with back pain are dissatisfied with medical care⁴ and seek care from complementary and alternative medical providers, including acupuncturists.^{5,6} Back pain is the leading reason for visits to licensed acupuncturists,⁷ and medical acupuncturists consider acupuncture an effective treatment for back pain.⁸

Several recent, well-designed European trials have suggested that real acupuncture and "sham" acupuncture (eg,

shallow needling of points considered ineffective) are equally effective^{9,10} and that both are superior to best-practice medical care,¹⁰ usual care,¹¹⁻¹³ and a wait-list control.⁹

Our trial expands on the findings of the European studies by (1) including a non-insertive method of stimulating acupuncture points, which permitted assessment of the need for needle insertion to achieve therapeutic benefit, (2) including both individualized and standardized forms of acupuncture, and (3) following patient outcomes for longer than most of the European trials. Thus, this trial was designed to address the following questions about the value of acupuncture for chronic low back pain:

1. Is acupuncture more effective than usual medical care alone?

2. Is real acupuncture more effective than simulated (noninsertive) acupuncture?
3. Is individualized acupuncture more effective than standardized acupuncture?

METHODS

STUDY DESIGN

We conducted a 4-arm randomized controlled trial comparing the effectiveness of individualized acupuncture, standardized acupuncture, simulated acupuncture, and usual care. Study design details are described elsewhere.¹⁴ This trial was approved by the institutional review boards of Group Health Cooperative, Seattle, Washington, and Kaiser Permanente of Northern California, Oakland. All participants gave written informed consent.

STUDY POPULATION

Patients aged 18 to 70 years who were receiving care for a back problem from an integrated health care delivery system in western Washington and another in northern California within the prior year were potentially eligible. We used electronic records to identify persons with diagnosis codes consistent with uncomplicated chronic low back pain within the prior 3 to 12 months. We excluded persons with (1) specific causes of back pain (eg, cancer, fractures, spinal stenosis, infections), (2) complicated back problems (eg, sciatica, prior back surgery, medicolegal issues), (3) possible contraindications for acupuncture (eg, coagulation disorders, cardiac pacemakers, pregnancy, seizure disorder), (4) conditions making treatment difficult (eg, paralysis, psychoses), and (5) conditions that might confound treatment effects or interpretation of results (eg, severe fibromyalgia, rheumatoid arthritis, concurrent care from other providers). Persons with less than 3 months of back pain or previous acupuncture treatment for any condition were excluded.

RECRUITMENT AND RANDOMIZATION PROCEDURES

Recruitment occurred from March 2004 through August 2006. Three to 12 months after back-related visits, potential participants were mailed invitation letters. Study staff telephoned respondents to determine final eligibility, which required a severity rating of at least 3 on the 0 to 10 back pain bothersomeness scale. We also mailed letters to members without recent visits for back pain and advertised in clinics and newsletters.

Those found eligible were administered a baseline questionnaire and randomly allocated to 1 of 4 treatment groups, using a centrally generated variable-sized block design. Treatments began within 2 weeks of randomization. The study was described only as a comparison of 3 methods of stimulating acupuncture points without information about how treatments differed.

STUDY TREATMENTS

Participants assigned to a real or simulated acupuncture treatment were treated twice weekly for 3 weeks and then weekly for 4 weeks (10 treatments total). Participants were asked to wear eye masks and lie prone with their heads in a face cradle. Electrostimulation, moxibustion, herbs, and other nonneedle adjuncts were proscribed.

One of 5 diagnostician acupuncturists with 7 to 18 years' experience evaluated participants at each visit using traditional Chinese medical diagnostic techniques and prescribed

individualized traditional Chinese medical treatments to be used only for participants randomized to individualized acupuncture. A therapist acupuncturist then delivered the assigned treatments, interacting minimally with participants and the diagnostician, who remained masked to treatment. Treatments were performed in research clinics at the 2 sites by 6 licensed acupuncturists with 4 to 19 years of experience. All acupuncturists were experienced in using traditional Chinese medical acupuncture for musculoskeletal pain. Of the 11 study acupuncturists, 9 had at least 3 years of formal training and the 2 others had practiced for over 15 years.

Acupuncturists used sterile disposable 32-gauge needles (0.25 mm) at least 1.5 inches in length. Needling depth varied slightly, depending on the acupuncture point, but was generally between 1 and 3 cm.

Individualized Acupuncture

This was the treatment prescribed by the diagnostician at the beginning of each visit. It could include any acupuncture points that could be needled with the participant lying prone. There were no constraints on number of needles, depth of insertion, or needle manipulation. Treatments averaged 10.8 needles (range, 5-20) retained for 18 minutes (range, 15-20 minutes). Seventy-four distinct points were used, half on the "Bladder meridian" that includes points on the back and legs.

Standardized Acupuncture

We used a standardized acupuncture prescription considered effective by experts for chronic low back pain.¹⁵ This included 8 acupuncture points commonly used for chronic low back pain (Du 3, Bladder 23—bilateral, low back ashi point, Bladder 40—bilateral, Kidney 3—bilateral) on the low back and lower leg.¹⁴ All acupuncture points were needled for 20 minutes, with stimulation by twirling the needles at 10 minutes and again just prior to needle removal. Therapists manipulated the needles to elicit "de qi," which they perceive as a biomechanical response in tissue as it tightens around the inserted needle and constricts its movement.¹⁶

Simulated Acupuncture

We developed a simulated acupuncture technique using a toothpick in a needle guide tube, which was found to be a credible acupuncture treatment by acupuncture-naïve patients with back pain.^{14,17} Simulating insertion involved holding the skin taut around each acupuncture point and placing a standard acupuncture needle guide tube containing a toothpick against the skin. The acupuncturist tapped the toothpick gently, twisting it slightly to simulate an acupuncture needle grabbing the skin, and then quickly withdrew the toothpick and guide tube while keeping his or her fingers against the skin for a few additional seconds to imitate the process of inserting the needle to the proper depth. All acupuncture points were stimulated with toothpicks at 10 minutes (ie, the acupuncturist touched each acupuncture point with the tip of a toothpick without the guide tube and rotated the toothpick clockwise and then counterclockwise less than 30°) and again at 20 minutes just before they were "removed." To simulate withdrawal of the needle, the acupuncturist tightly stretched the skin around each acupuncture point, pressed a cotton ball firmly on the stretched skin, then momentarily touched the skin with a toothpick (without the guide tube) and quickly pulled the toothpick away using the same hand movements as in regular needle withdrawal. The acupuncturists simulated insertion and removal of needles at the 8 acupuncture points used in the standardized treatment.

Usual Care Comparison Group

Participants in the usual care group received no study-related care—just the care, if any, they and their physicians chose (mostly medications, primary care, and physical therapy visits). All participants received a self-care book with information on managing flare-ups, exercise, and lifestyle modifications.¹⁸

OUTCOME MEASURES

Outcomes were measured at baseline and after 8, 26, and 52 weeks using computer-assisted telephone interviews by interviewers masked to treatment. Prespecified primary outcomes were back-related dysfunction and symptom bothersomeness at end of treatment (8 weeks). Dysfunction was measured using the modified Roland-Morris Disability Questionnaire (RMDQ), a reliable, valid, and sensitive measure¹⁹ appropriate for telephone administration. Participants were also asked to rate how bothersome their pain had been during the past week on a scale of 0 (“not at all bothersome”) to 10 (“extremely bothersome”). This measure demonstrates substantial construct validity.²⁰⁻²²

Secondary outcomes included the following: (1) 26- and 52-week outcomes for the primary outcome measures, (2) proportion of participants with clinically meaningful improvements²³ in dysfunction (≥ 3 point decrease on the RMDQ scale) and back pain (≥ 2 point decrease in symptom bothersomeness), (3) self-reported medication use for back pain in the prior week, (4) physical and mental health component summary scores of the Medical Outcomes Study Short-Form 36 Health Survey (SF-36),²⁴ and (5) number of days spent in bed, number of days lost from work or school, or cutting down on usual activities due to back problems during the past month.²⁵ Finally, participants' use of health services for back pain during the year following randomization was measured using interview data and, for the Washington site, automated health plan use data.

We also asked questions to determine if participants in the acupuncture groups perceived different experiences and to assess efforts to mask the diagnostician acupuncturists and outcomes assessors to study treatment.

Finally, participants were asked about adverse experiences at each visit and during the 8-week telephone follow-up assessment. This trial was monitored by the National Center for Complementary and Alternative Medicine Data Safety Monitoring Board.

STATISTICAL ANALYSES

Analyses were based on an intent-to-treat approach using randomized group assignment. The primary outcomes were analyzed as continuous measures. Analysis of covariance was used to test for treatment differences at the follow-up assessment, adjusting for the baseline measure. We also adjusted for site, age group (18-29, 30-39, 40-49, 50-59, and ≥ 60 years), and sex. Interactions of treatment group with age and site were added to test for effect modification by each of these covariates. Because expectations may influence outcomes, we also examined models that included expectation of acupuncture helpfulness and expectation of low back pain improvement as covariates and tested for effect modification with treatment groups dichotomized as treatment (real or simulated acupuncture) vs usual care.

Separate analyses were performed for each follow-up time. Results from the 8-week follow-up assessment were the primary end point. Pairwise comparisons using Tukey-Kramer adjustment were performed if the global test for differences among groups was significant at the $P < .05$ (2-sided) level. The study was powered to detect mean differences of 2.0 points on the RMDQ scale and 1.5 points on the symptom bothersomeness

scale, using variance estimates from our pilot study.¹⁴ Previous studies have suggested that these cutoff values are at the lower end of clinically important differences,²¹⁻²³ so their use should result in ample statistical power. We had 99% power to detect such differences in the overall analysis and approximately 80% power to detect pairwise differences after adjustment for multiple comparisons. To test for overall differences in our secondary outcomes, a χ^2 test was used for categorical outcomes and analysis of covariance was used for continuous outcomes. We used SAS/STAT statistical software (version 9.1; SAS institute Inc, Cary, NC).²⁶

RESULTS

STUDY RECRUITMENT AND FOLLOW-UP

We evaluated 2605 potential participants for eligibility; 641 (25%) were eligible and randomized (**Figure 1**). The main reasons for ineligibility were less than 3 months of back pain, sciatica, previous acupuncture, and inability to attend treatment visits. Three participants were excluded after randomization when we learned they had had exclusionary criteria when randomized (previous acupuncture treatment, involvement in litigation, and fibromyalgia). Therefore, analyses included 638 participants randomized to individualized acupuncture ($n = 157$), standardized acupuncture ($n = 158$), simulated acupuncture ($n = 162$), or usual care ($n = 161$). Follow-up rates were 95%, 91%, and 91% at 8, 26, and 52 weeks, respectively, and were similar across groups.

BASELINE CHARACTERISTICS

Study participants had a mean age of 47 years, and 62% were female, 68% were white, and 53% were college graduates (**Table 1**). Overall mean scores of 10.6 on the RMDQ scale and 5.1 for symptom bothersomeness indicated moderately severe chronic back problems. Two-thirds of participants reported at least 1 year of pain and current use of low back pain medication. Overall, participants were moderately optimistic that acupuncture would help (mean of 6.7 on a 0-10 scale).

STUDY TREATMENTS

A priori, we defined treatment adherence as completion of 8 or more of the 10 possible visits. By this definition, 84%, 87%, and 90% of participants were adherent in the individualized, standardized, and simulated acupuncture groups, respectively. At 8 weeks, 18% of participants reported having read more than two-thirds of the self-care book with no differences among groups ($P = .42$).

PRIMARY OUTCOMES

All groups showed improved function and decreased symptoms at the primary end point of 8 weeks (**Table 2** and **Figure 2**) in both the adjusted and unadjusted analyses. However, as seen in Table 2, unadjusted mean dysfunction scores for the individualized, standardized, and simulated acupuncture groups improved 4.4 to 4.5 points, compared with 2.1 points for those receiving usual care.

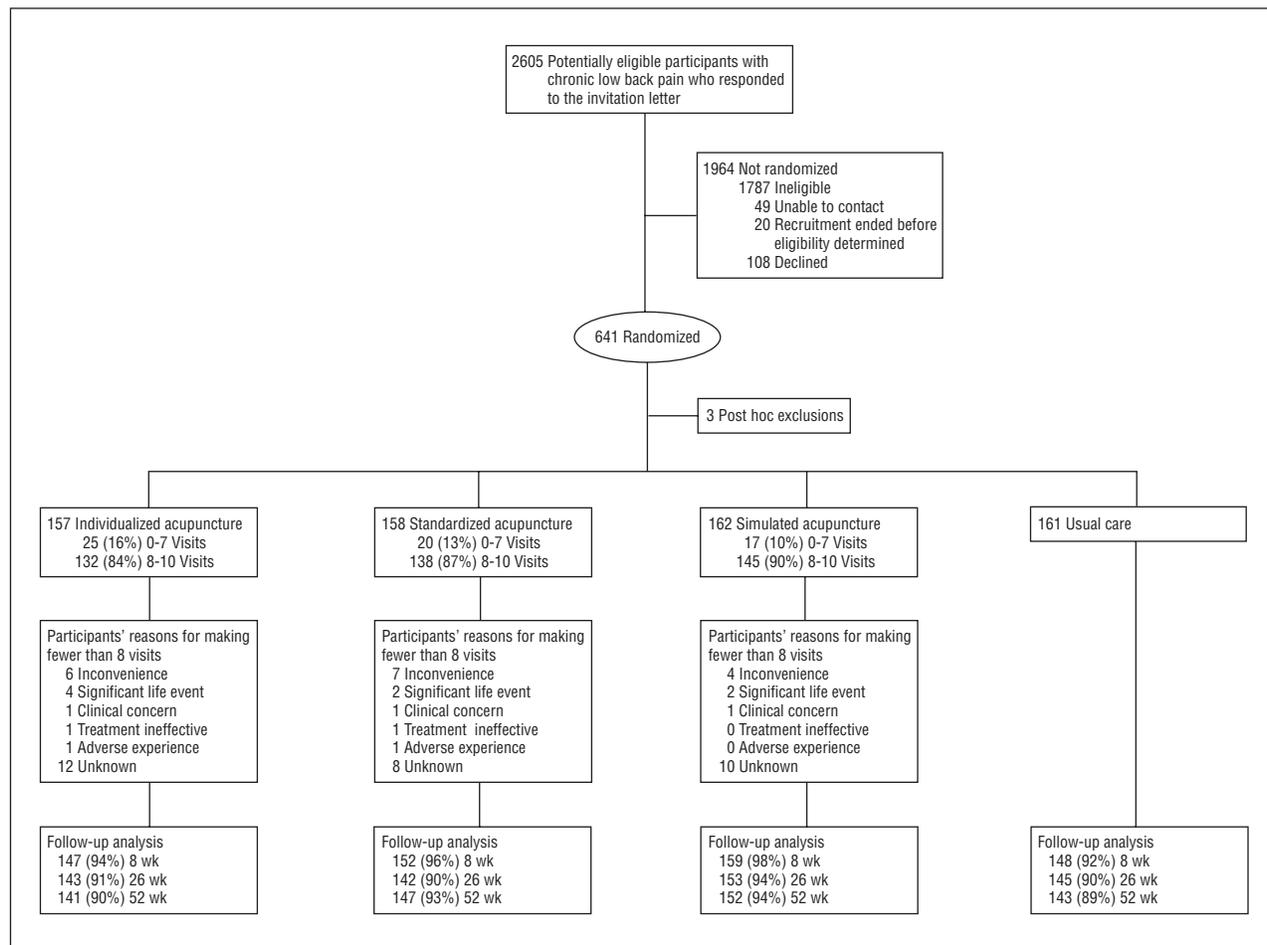


Figure 1. Study participant flow diagram.

There was a statistically significant difference in function among all 4 groups ($P < .001$ after adjustment for covariates) and statistically significant differences between usual care and each of the acupuncture groups adjusted for covariates and multiple comparisons (**Table 3**). However, there were no significant pairwise differences among the 3 acupuncture groups: individualized acupuncture was not significantly better than standardized acupuncture and real acupuncture was not significantly better than simulated acupuncture.

SECONDARY OUTCOMES

Mean values of the primary outcomes remained relatively stable from 8 to 52 weeks (Figure 2). The usual care group continued to have greater dysfunction than the real or simulated acupuncture groups through 52 weeks ($P = .001$). The real and simulated acupuncture groups did not differ significantly from one another, accounting for multiple comparisons ($P > .05$). The results for symptom bothersomeness were generally similar, but the differences among the 4 groups were smaller and no longer statistically significant at 52 weeks. Inclusion of the expectation measures did not alter the results and were not kept in the final models. There was no significant interaction between group and either age or site.

At 8 weeks, the proportion of participants improving at least 3 points on the RMDQ scale was about 60% in the real and simulated acupuncture groups, compared with only 39% in the usual care group (global test, $P < .001$) (**Figure 3A**). These superior outcomes in function for the real and simulated acupuncture groups remained significant at 26 weeks ($P = .01$) and 52 weeks ($P = .02$). Similar results were found for improvements of at least 2 points on the symptom bothersomeness score at 8 weeks ($P < .001$) (Figure 3B). However, overall differences were no longer significant at 26 or 52 weeks.

The use of medications for back pain in the past week (mostly nonsteroidal anti-inflammatory drugs) was similar across groups at baseline (ie, 62% to 65%), but by 8 weeks, it had decreased to 47% in the real and simulated acupuncture groups vs 59% in the usual care group ($P = .01$). This difference persisted at 26 and 52 weeks.

There was an overall group difference (favoring real and simulated acupuncture) at 8 weeks in both the SF-36 mental ($P = .03$) and physical ($P < .001$) component scores, but these differences were small (< 4 points) and no longer significant at 52 weeks. At 52 weeks, significantly more participants reported cutting down on activities for more than 7 days in the past month in the usual care group (18%) than in the real or simulated acupuncture groups (5%-7%) ($P < .001$). Similarly, more participants in the

Table 1. Baseline Characteristics of 638 Participants With Chronic Low Back Pain (LBP) by Treatment Group

Characteristic	Treatment Group				
	Individualized Acupuncture (n=157)	Standardized Acupuncture (n=158)	Simulated Acupuncture (n=162)	Usual Care (n=161)	Total (n=638)
Age, mean (SD), y	47 (13)	49 (13)	47 (14)	46 (13)	47 (13)
Female, %	68	56	60	64	62
White, %	67	69	72	65	68
Hispanic origin, %	7	9	8	8	8
College graduate, %	49	57	56	51	53
Married, %	55	60	61	60	59
Annual household income ≥\$45 000, %	71	71	67	75	71
Employed, %	78	80	78	81	79
Chronic pain for at least 1 y, %	69	74	60	70	68
Pain d in last 3 mo, mean (SD)	68 (26)	74 (22)	70 (24)	74 (23)	71 (24)
Pain below knee, %	21	22	21	21	21
Reduced activity ≥7 d in last 3 mo due to LBP, %	29	23	22	27	25
Reduced activity ≥7 d in last 4 wk due to LBP, %	27	25	20	28	25
Kept in bed or lying down ≥1 d in last 4 wk due to LBP, %	17	22	21	22	21
Kept from work or school for ≥1 d in last 4 wk due to LBP, %	19	22	22	18	21
SF-36 physical health scale score, mean (SD)	41 (9)	42 (8)	42 (8)	42 (8)	42 (8)
SF-36 mental health scale score, mean (SD)	53 (8)	54 (8)	54 (7)	53 (8)	53 (8)
Symptom bothersomeness in past wk (0-10 scale), mean (SD)	5.0 (2.4)	5.0 (2.3)	4.9 (2.3)	5.4 (2.3)	5.1 (2.3)
RMDQ score (0-23 scale), mean (SD)	10.8 (5.2)	10.8 (5.5)	9.8 (5.1)	11.0 (5.1)	10.6 (5.2)
Any medication use in past wk, %	62	62	63	65	63
Expect at least moderate improvement in LBP in 1 y, %	43	46	48	49	46
Expectation of acupuncture helpfulness (0-10 scale), mean (SD)	6.6 (1.8)	6.5 (1.8)	6.7 (1.9)	7.0 (1.8)	6.7 (1.8)

Abbreviations: RMDQ, Roland-Morris Disability Questionnaire; SF-36, Medical Outcomes Study Short-Form 36 Health Survey.

Table 2. Mean RMDQ Score and Symptom Bothersomeness Score at Baseline and Follow-up by Randomized Treatment Assignment

Score	Mean (SD)			
	Baseline	8 wk	26 wk	52 wk
RMDQ				
Individualized acupuncture	10.8 (5.2)	6.4 (5.3)	6.8 (5.5)	6.0 (5.4)
Standardized acupuncture	10.8 (5.6)	6.3 (5.7)	6.7 (5.8)	6.0 (5.8)
Simulated acupuncture	9.8 (5.2)	5.4 (4.9)	6.4 (6.0)	6.2 (5.8)
Usual care	11.0 (5.2)	8.9 (6.0)	8.4 (6.0)	7.9 (6.5)
Unadjusted <i>P</i> value		<.001	.01	.02
Adjusted <i>P</i> value ^a		<.001	.003	.001
Bothersomeness				
Individualized acupuncture	5.0 (2.5)	3.4 (2.7)	3.8 (2.5)	3.7 (2.6)
Standardized acupuncture	5.0 (2.3)	3.3 (2.5)	3.7 (2.6)	3.5 (2.7)
Simulated acupuncture	4.9 (2.4)	3.0 (2.4)	3.5 (2.7)	3.4 (2.7)
Usual care	5.4 (2.4)	4.7 (2.6)	4.4 (2.6)	4.1 (2.6)
Unadjusted <i>P</i> value		<.001	.03	.12
Adjusted <i>P</i> value ^a		<.001	.04	.12

Abbreviation: RMDQ, Roland-Morris Disability Questionnaire.

^aAdjusted for baseline outcome measure, site, age group (18-29, 30-39, 40-49, 50-59, 60-71 years), and sex.

usual care group missed work or school for more than a day in the past month (16%) than in the real or simulated acupuncture groups (5%-10%) ($P=.01$).

COINTERVENTIONS

Use of nonstudy treatments for back pain reported at the 8-week interview was similar across the real and simulated acupuncture groups, so pooled results are reported. Participants in the usual care group were twice as likely as those receiving real or simulated acupunc-

ture to report a physician or physical therapist visit (21% vs 11%; $P=.001$) or to have visited a complementary and alternative medicine provider (18% vs 8%; $P<.001$).

COSTS OF BACK-RELATED HEALTH CARE AFTER RANDOMIZATION

At the Washington site, mean total costs of back-related health services for the year after randomization were similar in the 4 treatment groups (range, \$160-\$221; $P=.65$). This excludes costs of the study's real and simulated acupunc-

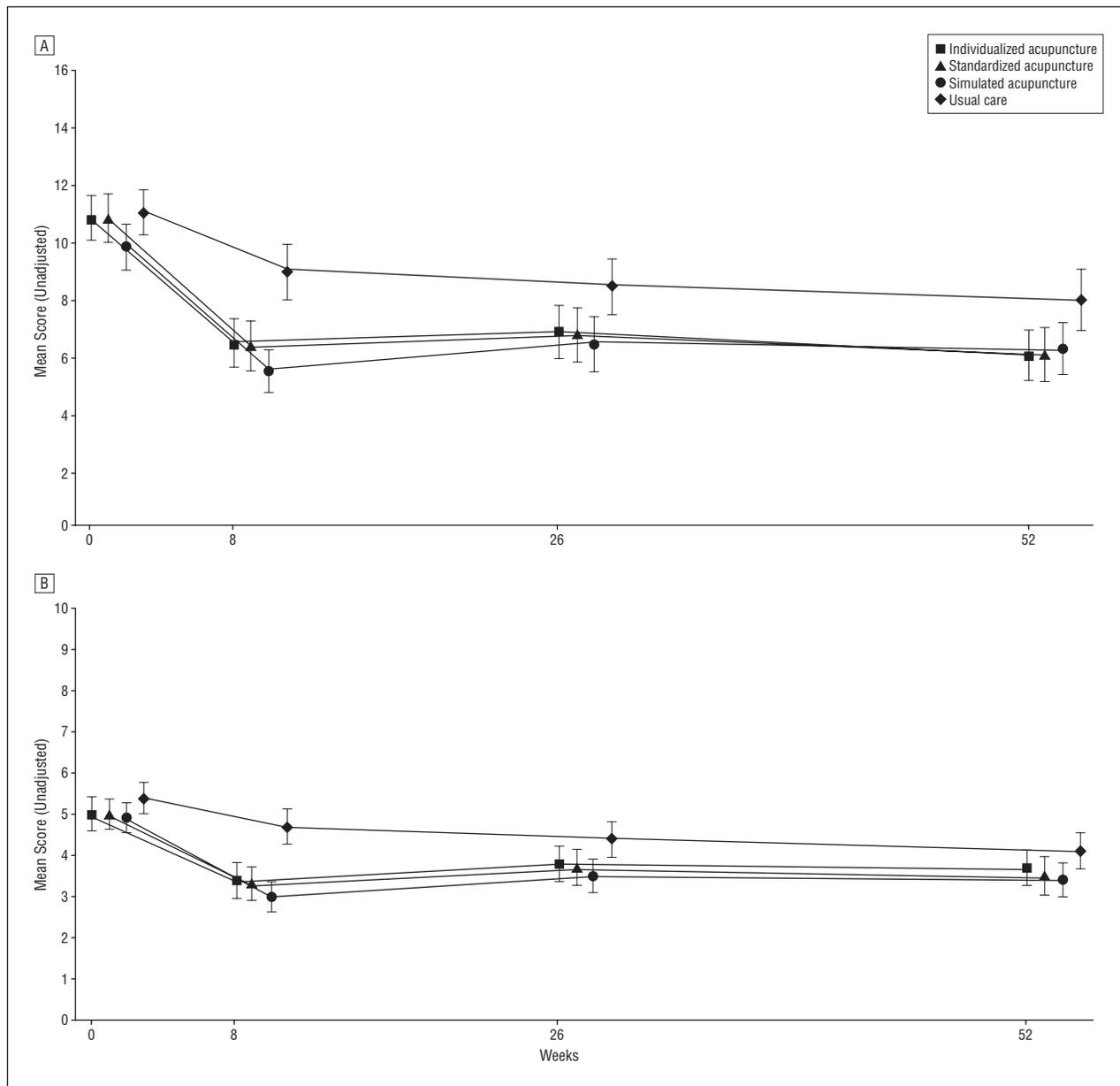


Figure 2. Mean Roland-Morris Disability Questionnaire scores (A) and symptom bothersomeness scores (B) and 95% confidence intervals by treatment group and time since randomization.

puncture treatments and the cost of the 1 spine operation in the usual care group.

TREATMENT CREDIBILITY AND MASKING

Participants rated the acupuncture and simulated acupuncture treatments almost identically with regard to provider skills and caring. The diagnostician acupuncturists rated the acupuncture and simulated acupuncture groups very similarly with regard to apparent efficacy and likelihood of receiving individualized treatment.

ADVERSE EVENTS

Of the 477 participants, 11 who were receiving real or simulated acupuncture reported a moderate adverse

experience possibly related to treatment (mostly short-term pain) and 1 reported a severe experience (pain lasting 1 month). One participant reported dizziness and another, back spasms. Rates of adverse experiences differed by treatment group: 6 of 157 participants for individualized acupuncture, 6 of 158 for standardized acupuncture, and 0 of 162 for simulated acupuncture ($P=.04$).

COMMENT

Compared with usual care, individualized acupuncture, standardized acupuncture, and simulated acupuncture had beneficial and persisting effects on chronic back pain. These treatments resulted in clinically meaningful improve-

Table 3. Adjusted^a Mean Difference and 95% Confidence Intervals for Pairwise Comparisons of RMDQ Score and Symptom Bothersomeness Score by Treatment Group and Time Since Randomization

Score	8 Weeks			26 Weeks			52 Weeks		
	Standardized Acupuncture	Simulated Acupuncture	Usual Care	Standardized Acupuncture	Simulated Acupuncture	Usual Care	Standardized Acupuncture	Simulated Acupuncture	Usual Care
RMDQ									
Individualized acupuncture	0.16 (-0.90 to 1.22)	0.45 (-0.61 to 1.50)	-2.47 ^b (-3.53 to -1.40)	-0.05 (-1.18 to 1.08)	-0.41 (-1.53 to 0.70)	-1.86 ^b (-2.98 to -0.74)	-0.26 (-1.40 to 0.87)	-1.10 (-2.23 to 0.04)	-2.08 ^b (-3.22 to -0.94)
Standardized acupuncture		0.29 (-0.76 to 1.33)	-2.63 ^b (-3.69 to -1.56)		-0.36 (-1.48 to 0.75)	-1.81 ^b (-2.94 to -0.68)		-0.84 (-1.95 to 0.28)	-1.82 ^b (-2.95 to -0.69)
Simulated acupuncture			-2.91 ^b (-3.96 to -1.86)			-1.45 (-2.56 to -0.34)			-0.98 (-2.11 to 0.14)
Bothersomeness									
Individualized acupuncture	0.22 (-0.34 to 0.77)	0.52 (-0.03 to 1.07)	-1.05 ^b (-1.60 to -0.49)	0.17 (-0.43 to 0.77)	0.24 (-0.35 to 0.83)	-0.54 (-1.13 to 0.06)	0.20 (-0.39 to 0.79)	0.17 (-0.42 to 0.76)	-0.45 (-1.04 to 0.15)
Standardized acupuncture		0.30 (-0.24 to 0.85)	-1.26 ^b (-1.82 to -0.71)		0.07 (-0.52 to 0.66)	-0.71 (-1.31 to -0.11)		-0.03 (-0.61 to 0.55)	-0.65 (-1.24 to -0.06)
Simulated acupuncture			-1.56 ^b (-2.11 to -1.02)			-0.78 ^b (-1.36 to -0.19)			-0.62 (-1.21 to -0.03)

Abbreviation: RMDQ, Roland-Morris Disability Questionnaire.

^aAdjusted for baseline outcome measure, site, age group (18-29, 30-39, 40-49, 50-59, 60-71 years), and sex.

^bSignificant at the $P < .05$ level after adjustment for multiple comparisons.

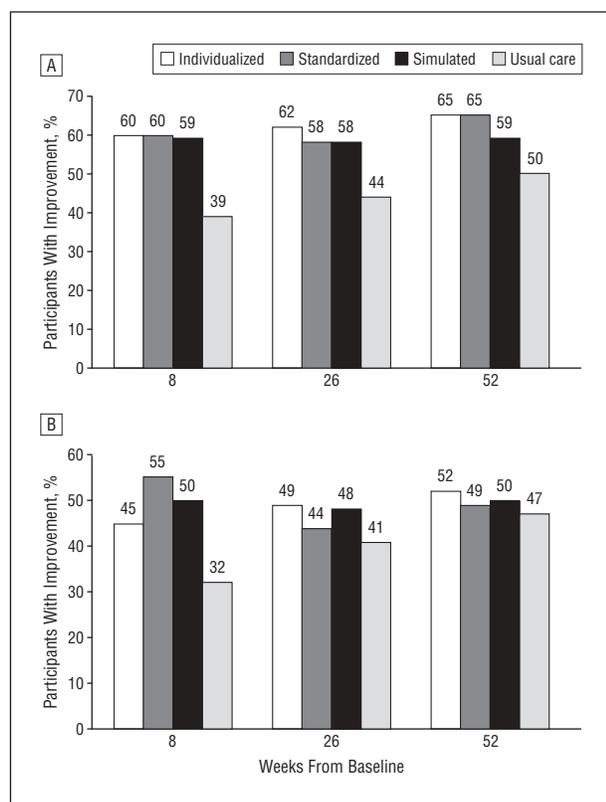


Figure 3. Participants with improvement. Percentage of participants improving by at least 3 points on the Roland-Morris Disability Questionnaire scale (A) and by at least 2 points on the symptom bothersomeness scale (B).

ments in function. Substantial adverse experiences with needle insertion were infrequent (1 of 315 participants). Self-reported medication use in the real and simulated acupuncture groups decreased significantly more than in the usual care group and remained lower through the 1-year follow-up. However, the 8 to 10 acupuncture treatments received by most participants (which would cost between \$600 and \$1200) did not result in cost savings to the health plan during the year after randomization.

This trial differs from our earlier study (which found similar effects for acupuncture and a more rigorous educational intervention) by including a usual care group, participants with more chronic pain, and participants who were all acupuncture naïve.²⁷ However, our findings are consistent with those of recent high-quality trials. One German trial found that both real acupuncture and sham acupuncture (superficial needling at nonacupuncture points) had similar effects that were superior to those of guideline-based conventional medical treatment.¹⁰ A second German trial found that both real and sham acupuncture were superior to a wait list control group but not significantly different from each other.⁹ Finally, a British trial found that traditional acupuncture care delivered in a primary care setting had modestly superior results compared with usual care after 2 years.¹² Our trial extends the findings from these studies by demonstrating that needle insertion is not necessary to achieve therapeutic benefits and by measuring longer-term outcomes.

Collectively, these recent trials provide strong and consistent evidence that real acupuncture needling using the Chinese meridian system is no more effective for chronic back pain than various purported forms of sham acupuncture. However, both real and sham acupuncture appear superior to usual care. Possible explanations for these findings include the following: (1) superficial acupuncture point stimulation directly stimulates physiological processes that ultimately lead to improved pain and function, or (2) participants' improved functioning resulted from nonspecific effects such as therapist conviction, patient enthusiasm, or receiving a treatment believed to be helpful.

The appropriateness of using minimal, superficial, or sham control groups in trials of acupuncture remains controversial.²⁸ In fact, the use of blunt needles that did not penetrate the skin was described 2000 years ago in the classic book on acupuncture needling.²⁹ A study using functional magnetic resonance imaging found that superficial and deep needling of an acupuncture point elicited similar blood oxygen level-dependent responses.³⁰ Another study demonstrated that lightly touching the skin can stimu-

late mechanoreceptors that induce emotional and hormonal reactions, which in turn alleviate the affective component of pain.³¹ This could explain why trials evaluating acupuncture for pain have failed to find that real acupuncture is superior to sham or superficial control treatments and raises questions about whether sham treatments truly serve as inactive controls.

The possibility that an acupuncture treatment “experience” could be beneficial because of nonspecific effects is also credible.³² A recent acupuncture trial for irritable bowel syndrome reported that nonspecific effects (especially the patient-clinician relationship) produced statistically and clinically significant outcomes.^{33,34} The potency of nonspecific effects has also been noted in placebo-controlled randomized trials of surgical interventions for pain conditions.^{35,36}

The main strengths of this trial are its size, high compliance and follow-up rates, long-term follow-up, inclusion of a simulated acupuncture control, and effective masking. Limitations include restricting treatment to a single component (needling) of normal traditional Chinese medicine acupuncture,³⁷ prespecification of the number and duration of treatments, limited conversation between acupuncturists and participants, and exclusion of a medical attention control group. However, a recent trial using a similar number and duration of visits for both the acupuncture and medical care control groups also found that acupuncture was superior.¹⁰

Our results have important implications for key stakeholders. For clinicians and patients seeking a relatively safe^{38,39} and effective treatment for a condition for which conventional treatments are often ineffective, various methods of acupuncture point stimulation appear to be reasonable options, even though the mechanism of action remains unclear. Furthermore, the reduction in long-term exposure to the potential adverse effects of medications is an important benefit that may enhance the safety of conventional medical care. The number of patients who would need to be treated with insertive or superficial acupuncture stimulation to result in 1 person achieving meaningful improvement in function ranges from 5 (for short-term benefits) to 8 (for persisting benefits).

In conclusion, acupuncturelike treatments significantly improved function in persons with chronic low back pain. However, the finding that benefits of real acupuncture needling were no greater than those of noninsertive stimulation raises questions about acupuncture’s purported mechanism of action. Future research is needed to determine the relative contributions of the physiologic effects of noninsertive stimulation, patient expectations, and other nonspecific effects.

Accepted for Publication: December 17, 2008.

Correspondence: Daniel C. Cherkin, PhD, Center for Health Studies, 1730 Minor Ave, Ste 1600, Seattle, WA 98101 (cherkin.d@ghc.org).

Author Contributions: The principal investigator, Dr Cherkin, had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. *Study concept and design:* Cherkin, Sherman, Barlow, and Deyo. *Acquisition of data:* Cherkin, Sherman, Avins, Erro, Barlow, Delaney,

Hawkes, Hamilton, and Pressman. *Analysis and interpretation of data:* Cherkin, Sherman, Avins, Ichikawa, Barlow, Delaney, Hamilton, Pressman, Khalsa, and Deyo. *Drafting of the manuscript:* Cherkin, Sherman, Ichikawa, and Barlow. *Critical revision of the manuscript for important intellectual content:* Cherkin, Sherman, Avins, Erro, Ichikawa, Barlow, Delaney, Hawkes, Hamilton, Pressman, Khalsa, and Deyo. *Statistical analysis:* Sherman, Ichikawa, Barlow, Delaney, and Pressman. *Obtained funding:* Cherkin and Avins. *Administrative, technical, and material support:* Cherkin, Avins, Erro, Delaney, Hawkes, Hamilton, Pressman, Khalsa, and Deyo. *Study supervision:* Cherkin, Sherman, Erro, Barlow, Hawkes, and Hamilton.

Financial Disclosure: None reported.

Funding/Support: This trial was funded through a National Institutes of Health (NIH) Cooperative Agreement (U01 AT 001110) with the National Center for Complementary and Alternative Medicine (NCCAM). The sponsor (NIH), through its project officer, Dr Khalsa, was involved in the analysis and interpretation of data and review and approval of the manuscript. Lhasa OMS Inc, Weymouth, Massachusetts, donated the Seirin acupuncture needles used in this study.

Disclaimer: The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of NCCAM.

Previous Presentation: Portions of the results described in the article were presented at the the Ninth International Forum for Primary Care Research on Low Back Pain; October 6, 2007; Majorca, Spain.

Additional Contributions: The following people provided assistance with this study: Sara Bayer, LAc, Ramey Fair, LAc, Larry Forsberg, LAc, Roxanne Geller, LAc, Caryn Goldman, LAc, Paul Griffin, LAc, Susan M. Kaetz, MPH, LAc, Ken Morris, LAc, Sabi Inderkum, LAc, Sachiko Nakano, LAc, Deborah Stanfill, LAc, and Eliot Wagner, LAc (diagnostician and therapist acupuncturists in private practice); Zoe Bermet, Marissa Brooks, John Ewing, Erika Holden, Danielle Huston, Christel Kratochvil, and Melissa Parson (clinic research assistants at Group Health Center for Health Studies); Olivia Anaya, Pete Bogdanos, Cynthia H. Huynh, Rebecca Rogot, and Caroline M. Sison (clinic research assistants at Kaiser Permanente Northern California Division of Research); Kabba Anand, DAC, LAc (consultant acupuncturist in private practice); Harley Goldberg, DO (physician with Kaiser Permanente Northern California); Juanita Jackson (administrative assistant, Group Health Center for Health Studies); and John Maio (programmer, Kaiser Permanente Northern California Division of Research). All of these individuals were compensated for their roles in the project. Our original NCCAM project officer, Richard Nahin, PhD, MPH, provided helpful advice.

REFERENCES

1. Luo X, Pietrobon R, Sun SX, Liu GG, Hey L. Estimates and patterns of direct health care expenditures among individuals with back pain in the United States. *Spine*. 2004;29(1):79-86.
2. Martin BI, Deyo RA, Mirza SK, et al. Expenditures and health status among adults with back and neck problems. *JAMA*. 2008;299(6):656-664.
3. Stewart WF, Ricci JA, Chee E, Morganstein D, Lipton R. Lost productive time

- and cost due to common pain conditions in the US workforce. *JAMA*. 2003; 290(18):2443-2454.
4. Roter D, Hall J. How is your doctor treating you? *Consumer Reports*. February 1995;60:81-88.
 5. Wolsko PM, Eisenberg DM, Davis RB, Kessler R, Phillips RS. Patterns and perceptions of care for treatment of back and neck pain: results of a national survey. *Spine*. 2003;28(3):292-298.
 6. Barnes PM, Powell-Griner E, McFann K, Nahin RL. Complementary and alternative medicine use among adults: United States, 2002. *Adv Data*. May 27, 2004; (343):1-19.
 7. Cherkin DC, Deyo RA, Sherman KJ, et al. Characteristics of visits to licensed acupuncturists, chiropractors, massage therapists, and naturopathic physicians. *J Am Board Fam Pract*. 2002;15(6):463-472.
 8. Diehl DL, Kaplan G, Coulter I, Glik D, Hurwitz EL. Use of acupuncture by American physicians. *J Altern Complement Med*. 1997;3(2):119-126.
 9. Brinkhaus B, Witt CM, Jena S, et al. Acupuncture in patients with chronic low back pain: a randomized controlled trial. *Arch Intern Med*. 2006;166(4):450-457.
 10. Haake M, Muller HH, Schade-Brittinger C, et al. German Acupuncture Trials (GERAC) for chronic low back pain: randomized, multicenter, blinded, parallel-group trial with 3 groups. *Arch Intern Med*. 2007;167(17):1892-1898.
 11. Thomas KJ, MacPherson H, Ratcliffe J, et al. Longer term clinical and economic benefits of offering acupuncture care to patients with chronic low back pain. *Health Technol Assess*. 2005;9(32):iii-iv, ix-x, 1-109.
 12. Thomas KJ, MacPherson H, Thorpe L, et al. Randomised controlled trial of a short course of traditional acupuncture compared with usual care for persistent non-specific low back pain. *BMJ*. 2006;333(7569):623.
 13. Witt CM, Jena S, Selim D, et al. Pragmatic randomized trial evaluating the clinical and economic effectiveness of acupuncture for chronic low back pain. *Am J Epidemiol*. 2006;164(5):487-496.
 14. Cherkin DC, Sherman KJ, Hogeboom CJ, et al. Efficacy of acupuncture for chronic low back pain: protocol for a randomized controlled trial. *Trials*. February 2008; 9:10.
 15. Sherman KJ, Cherkin DC. Developing methods for acupuncture research: rationale for and design of a pilot study evaluating the efficacy of acupuncture for chronic low back pain. *Altern Ther Health Med*. 2003;9(5):54-60.
 16. Langevin HM, Churchill DL, Fox JR, Badger GJ, Garra BS, Krag MH. Biomechanical response to acupuncture needling in humans. *J Appl Physiol*. 2001;91(6): 2471-2478.
 17. Sherman KJ, Hogeboom CJ, Cherkin DC, Deyo RA. Description and validation of a noninvasive placebo acupuncture procedure. *J Altern Complement Med*. 2002; 8(1):11-19.
 18. Moore JE, Lorig K, Von Korff M, Gonzalez V, Laurent D. *The Back Pain Helpbook*. Reading, MA: Perseus Publishing; 1999.
 19. Bombardier C. Outcome assessments in the evaluation of treatment of spinal disorders: summary and general recommendations. *Spine*. 2000;25(24):3100-3103.
 20. Dunn KM, Croft PR. Classification of low back pain in primary care: using "bothersomeness" to identify the most severe cases. *Spine*. 2005;30(16): 1887-1892.
 21. Patrick DL, Deyo RA, Atlas SJ, Singer DE, Chapin A, Keller RB. Assessing health-related quality of life in patients with sciatica. *Spine*. 1995;20(17):1899-1909.
 22. Cherkin DC, Deyo RA, Street JH, Barlow W. Predicting poor outcomes for back pain seen in primary care using patients' own criteria. *Spine*. 1996;21(24): 2900-2907.
 23. Ostelo RW, Deyo RA, Stratford P, et al. Interpreting change scores for pain and functional status in low back pain: towards international consensus regarding minimal important change. *Spine*. 2008;33(1):90-94.
 24. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36), I: conceptual framework and item selection. *Med Care*. 1992;30(6):473-483.
 25. Current estimates from the national health interview survey: United States, 1984. *Vital Health Stat 10*. July 1986;(156):1-191.
 26. *SAS/STAT 9.1 User's Guide, Version 8*. Cary, NC: SAS Institute Inc; 2004.
 27. Cherkin DC, Eisenberg D, Sherman KJ, et al. Randomized trial comparing traditional Chinese medical acupuncture, therapeutic massage, and self-care education for chronic low back pain. *Arch Intern Med*. 2001;161(8):1081-1088.
 28. Lewith GT, White PJ, Kaptchuk TJ. Developing a research strategy for acupuncture. *Clin J Pain*. 2006;22(7):632-638.
 29. Birch SJ, Felt RL. *Understanding Acupuncture*. New York, NY: Churchill Livingstone; 1999.
 30. MacPherson H, Green G, Nevado A, et al. Brain imaging of acupuncture: comparing superficial with deep needling. *Neurosci Lett*. 2008;434(1):144-149.
 31. Lund I, Lundeberg T. Are minimal, superficial or sham acupuncture procedures acceptable as inert placebo controls? *Acupunct Med*. 2006;24(1):13-15.
 32. Di Blasi Z, Harkness E, Ernst E, Georgiou A, Kleijnen J. Influence of context effects on health outcomes: a systematic review. *Lancet*. 2001;357(9258):757-762.
 33. Kaptchuk TJ, Kelley JM, Conboy LA, et al. Components of placebo effect: randomized controlled trial in patients with irritable bowel syndrome. *BMJ*. 2008; 336(7651):999-1003.
 34. Spiegel D, Harrington A. What is the placebo worth? *BMJ*. 2008;336(7651):967-968.
 35. Cobb LA, Thomas GI, Dillard DH, Merendino KA, Bruce RA. An evaluation of internal-mammary-artery ligation by a double-blind technic. *N Engl J Med*. 1959;260 (22):1115-1118.
 36. Moseley JB, O'Malley K, Petersen NJ, et al. A controlled trial of arthroscopic surgery for osteoarthritis of the knee. *N Engl J Med*. 2002;347(2):81-88.
 37. Sherman KJ, Cherkin DC, Eisenberg DM, Erro J, Hrbek A, Deyo RA. The practice of acupuncture: who are the providers and what do they do? *Ann Fam Med*. 2005; 3(2):151-158.
 38. MacPherson H, Thomas K, Walters S, Fitter M. A prospective survey of adverse events and treatment reactions following 34,000 consultations with professional acupuncturists. *Acupunct Med*. 2001;19(2):93-102.
 39. 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(7311):485-486.