Acupuncture versus Placebo for the Treatment of Chronic Mechanical Neck Pain

A Randomized, Controlled Trial

Peter White, PhD, BSc; George Lewith, DM, FRCP; Phil Prescott, PhD, DIC, ARCS, BSc; and Joy Conway, PhD

Background: **Despite substantial increases in its popularity and use, the efficacy of acupuncture for chronic mechanical neck pain remains unproved.**

Objective: **To compare acupuncture and placebo for neck pain.**

Design: **A randomized, single-blind, placebo-controlled, parallel-arm trial with 1-year follow-up.**

Setting: **The outpatient departments of 2 major hospitals in the United Kingdom, 1999 to 2001.**

Patients: **135 patients 18 to 80 years of age who had chronic mechanical neck pain. Eleven patients withdrew from treatment, and 124 completed the primary end point.**

Measurements: **The primary outcome was pain 1 week after treatment, according to a visual analogue scale. Secondary outcomes were pain at other time points, score on the Neck Disability Index and the Short Form-36, and use of analgesic medications.**

Interventions: **Patients were randomly assigned to receive, over 4 weeks, 8 treatments with acupuncture or with mock transcutaneous electrical stimulation of acupuncture points using a decommissioned electroacupuncture stimulation unit.**

Results: **Both groups improved statistically from baseline, and acupuncture and placebo had similar credibility. For the primary outcome (weeks 1 to 5), a statistically significant difference in visual analogue scale score in favor of acupuncture (6.3 mm [95% CI, 1.4 to 11.3 mm]; $P = 0.01$) was observed between the 2 study groups, after adjustment for baseline pain and other covariates. However, this difference was not clinically significant because it demonstrated only a 12% (CI, 3% to 21%) difference between acupuncture and placebo. Secondary outcomes showed a similar pattern.**

Limitations: **All treatments were provided by 1 practitioner. Although the control was credible, it did not mimic the process of needling. A nonintervention group was not present to control for regression to the mean.**

Conclusions: **Acupuncture reduced neck pain and produced a statistically, but not clinically, significant effect compared with placebo. The beneficial effects of acupuncture for pain may be due to both nonspecific and specific effects.**

Ann Intern Med. 2004;141:911-919. www.annals.org

For author affiliations, see end of text.

See related articles on pp 901-910 and pp 920-928 and editorial comment on pp 957-958.

Chronic mechanical neck pain can be caused by dysfunction of a variety of structures within the neck (1–3) but specifically excludes systemic problems such as rheumatoid arthritis. It is usually associated with unspecified degenerative changes (cervical spondylosis) that include osteoarthritis. Neck pain presents a substantial problem and may be responsible for as many days of work absenteeism as low back pain (4). Osteoarthritis is the most common of the chronic diseases and affects most people older than 65 years of age (5). It is degenerative and progressive in nature (1–3, 6). Because there is no cure for cervical spondylosis, treatment tends to center on symptom relief (7, 8). If the condition is symptomatic and symptoms are left untreated, this manifests as increasing episodic pain, stiffness, or both, and patients may then experience a spiral of increasing dysfunction (9).

Two systematic reviews of conventional conservative physical therapy suggest that little evidence supports the efficacy of such intervention (8, 10). Nonsteroidal anti-inflammatory drugs, the mainstay of pharmacologic treatment, are associated with a wide spectrum of serious and well-documented adverse reactions (5). There has been a huge increase in the use of complementary and alternative medicine in both the United States and the United Kingdom (11, 12). Acupuncture is the most frequently used type of complementary and alternative medicine therapy for the treatment of osteoarthritis (13). Despite little sound evidence of efficacy (14), approximately 1 million people seek complementary and alternative medicine treatment annually in the United States (15). Two systematic reviews of acupuncture in neck pain (4, 16) suggest that there has been insufficient research in this area and that the published studies are of poor quality. Therefore, we conducted a rigorous evaluation to test whether western-style acupuncture performs better than placebo for treatment of chronic mechanical neck pain. We defined *western acupuncture* as a conventional diagnosis followed by individualized acupuncture treatment using a combination of prescriptive tender, local, and distal points. This is in contrast to a traditional Chinese approach, which would formulate an individualized diagnosis based on traditional Chinese theories of meridians and energy (or *qi*).

METHODS

Study Design and Patient Selection

We designed a randomized, single-blind, placebo-controlled trial using a pragmatic treatment regimen. The study was conducted in the outpatient departments of Southampton General Hospital and Salisbury District

Context

Chronic neck pain is a common, disabling condition and is difficult to treat. Because traditional therapies, such as physical therapy and nonsteroidal anti-inflammatory drugs, have limited effectiveness and can have adverse effects, many patients seek alternative therapies, such as acupuncture.

Contribution

This randomized, controlled trial of acupuncture versus sham transcutaneous electrical stimulation for patients with chronic mechanical neck pain identified no clinically significant benefit of acupuncture over placebo with respect to pain, function, or analgesic use.

Implications

Acupuncture did not lead to clinically significant improvement in chronic mechanical neck pain when compared with placebo.

–The Editors

Hospital in the United Kingdom. Between 1999 and 2001, patients were referred by rheumatologists or family physicians or from physiotherapy waiting lists. Appropriate ethical approval was obtained from the Southampton and South West Hampshire Joint Research Ethics Committee. Patients were 18 to 80 years of age, had chronic (>2 months) mechanical neck pain, and had a pain score of more than 30 mm on a visual analogue scale (VAS) for 5 of 7 pretreatment days (possible score on this VAS ranges from 0 to 100 mm). We excluded pregnant patients; those with a history of fracture or surgery to the neck, cervical congenital abnormality, uncontrolled low back pain, contraindication to acetaminophen, systemic illness (for example, rheumatoid arthritis), or ongoing neck-related litigation or disability claims; and those with current or recent manual neck treatment or steroid use (oral or local injection).

Sample Size

Previous studies suggested a 50% to 75% improvement using acupuncture for patients with chronic pain, compared with 30% for placebo treatment (17). We used these findings as the basis for our power calculation. At 5% significance and 90% power, 53 patients were required in each group to detect a 30% difference in response between active and placebo treatments.

Interventions

Informed consent was obtained along with a full medical history, neurologic examination, cervical radiography, and laboratory investigations (full blood count, erythrocyte sedimentation rate, and liver function tests). We told patients that we were comparing 2 types of intervention: 1) acupuncture with needles and 2) treatment with a machine designed to stimulate acupuncture points through skin electrodes. The patients were informed that the treatment might or might not prove to be effective and that there was a 50% chance that they would be assigned to a placebo but were not told what the placebo was. Blinding was not broken until 1 year after treatment. One practitioner performed all interventions.

Acupuncture

We used single-use, sterile, silver-handle, prepacked needles without guide tubes. Sizes used were 13 mm × 0.25 mm, 25 mm × 0.25 mm, and 40 mm × 0.25 mm. We based point selection on individualized western acupuncture techniques by using a list of points previously reported as being effective in neck pain (18, 19) and by reaching a consensus according to our own clinical and teaching practice (**Appendix Table**, available at www.annals.org). The specific points for each individual were defined at each treatment session, depending on the patient's pain distribution and palpation of the neck and thorax to determine *ah-shi* points, or local tender points, for acupuncture. At least 1 distal point was used. Point location and depth of insertion were as described in traditional texts (19).

Six points on average, per side if pain was bilateral, were used on each patient, and *deqi* (a term used to describe acupuncture needle sensation) was obtained on each needle. Twenty-minute treatment sessions were given. The patient was checked every 6 or 7 minutes to ascertain whether *deqi* was still present, and needles were manipulated again if required.

Placebo

There is considerable debate about the ideal placebo for acupuncture studies (20). To enhance the rigor of our study, we chose to use a previously well-validated placebo (17, 21–23) that could not have a specific physiologic effect. The Noma FM-4 electroacupuncture stimulator (Noma Ltd., Southampton, United Kingdom) was used. It has 4 channels, allowing pseudostimulation of up to 8 acupuncture points simultaneously, and emits visual and audio signals. Reusable electrodes (Body Clock Health Care Ltd., London, United Kingdom) were fixed to the surface of the patient's skin and were connected to the stimulator through decommissioned cables. The cables were severed inside the output plug, so that no current could reach the patient. Examination and point selection were the same as with real acupuncture for each treatment. Point location and treatment variables were changed during subsequent treatment sessions if patients felt they were not progressing. Patients were told that the machine could stimulate acupuncture points through high-frequency, low-intensity stimulation and therefore would not produce any sensation. If patients reported sensation, the therapist adjusted the unit for comfort (although since this was a sham procedure, such adjustment made no real difference).

Patients in both groups were instructed to use acetaminophen alone for pain relief and were not given or permitted any other form of treatment, including exercises or stretches, during the study and for 2 months after treatment ended. Patients were treated twice per week for 4 weeks. The therapist had taken an Acupuncture Association of Chartered Physiotherapists accredited course on western acupuncture techniques and had 7 years of experience practicing acupuncture.

Randomization

Randomization lists for Southampton General Hospital and Salisbury District Hospital were generated by using a computer program, Randomlogue, produced by Southampton University Department of Medical Statistics, Southampton, United Kingdom. Sealed envelopes containing the individual randomization codes, numbered consecutively with sex and age strata, were then prepared (sex strata were male and female, and age strata were 18 to 49 years and 50 to 80 years) Patients were strictly allocated across 4 strata for either of the study hospitals according to the next available envelope number.

Data Collection

Before randomization, patients were instructed to use a daily pain diary (on a VAS) to record pain and acetaminophen use for 7 consecutive days. After a week, the pain diaries were examined, and if the inclusion criteria (which were unknown to the patients) were satisfied, randomization occurred. Patients completed the various questionnaires before treatment and at 1 and 8 weeks after treatment. Pain continued to be recorded either weekly or daily throughout treatment and up to 8 weeks after treatment. Additional questionnaires and pain evaluation were also completed at 6 and 12 months after treatment. **Figure 1** shows a flow chart of the trial.

Outcomes

The VAS score for pain 1 week after treatment, recorded daily for 7 consecutive days, was the primary outcome. Pain was recorded on a scale of 0 mm (indicating no pain) to 100 mm (indicating worst pain imaginable). A significant outcome for acupuncture was considered to be a 30% difference in VAS score between groups 5 weeks after randomization (1 week after treatment), taking into account baseline pain.

Secondary Outcomes

Pain (as measured by VAS score) was also used as a secondary outcome at other time points. When pain was recorded on a weekly basis, patients were asked, "What was your average daily pain over the last seven days?" Patients completed the Neck Disability Index (NDI), a previously validated disease-specific quality-of-life measurement (24–28), at various time points (**Figure 1**). Lower scores indicated better health. We also used the Short Form-36 (SF-36), a quality-of-life measure that has been shown to have good reliability, sensitivity, and validity (24, 29–33) (**Figure 1**). The SF-36 gives 2 scores, a Physical Component Summary score and a Mental Component Summary score. Higher scores indicate better health. Patients also recorded acetaminophen use in their diaries.

A Borkovec and Nau scale, which has previously been used for acupuncture studies (34, 35), was used to ascertain treatment credibility (patient equipoise) in the 2 groups. Before treatment, patients were asked, "How confident do you feel that this treatment can alleviate your complaint?" and "How confident would you be in recommending this treatment to a friend with a similar complaint?" After treatment, patients were asked, "How logical does this treatment seem to you?" and "How successful do you think this treatment would be in alleviating other complaints?" Questions were scored on a 0 to 6 Likert scale, with 6 indicating higher credibility.

Statistical Analysis

The mean daily VAS score was determined at baseline, at week 5 (1 week post-treatment), and at week 12 (8 weeks post-treatment) for each group. Weekly VAS score was used for the intermediate weeks. The VAS score over the pretreatment week was compared with the VAS score at other time points by using a multivariate longitudinal analysis with generalized estimating equations, adjusting for baseline pain and for other confounders (age, sex, hospital, length of illness, and analgesic use). We performed the longitudinal analyses by using Stata (Stata Corp., College Station, Texas) on an intention-to-treat basis over the 5-week and 12-week periods. Generalized estimating equations were used in the longitudinal analysis, with an autoregressive correlation structure between the weekly scores for the same patient (36–38). This type of data analysis was a departure from our original protocol, which was finalized in 1999. We originally intended to use analysis of covariance as the primary statistical method but introduced longitudinal analysis after peer review. Intake of analgesic medications was used as a potential confounding factor in the longitudinal analysis.

Changes in the secondary variables, NDI and SF-36 scores, were also examined by using a multivariate longitudinal analysis with adjustments for baseline and other covariates. Frequencies of responses to the credibility questions were compared for each group by using the chi-square test.

Role of the Funding Source

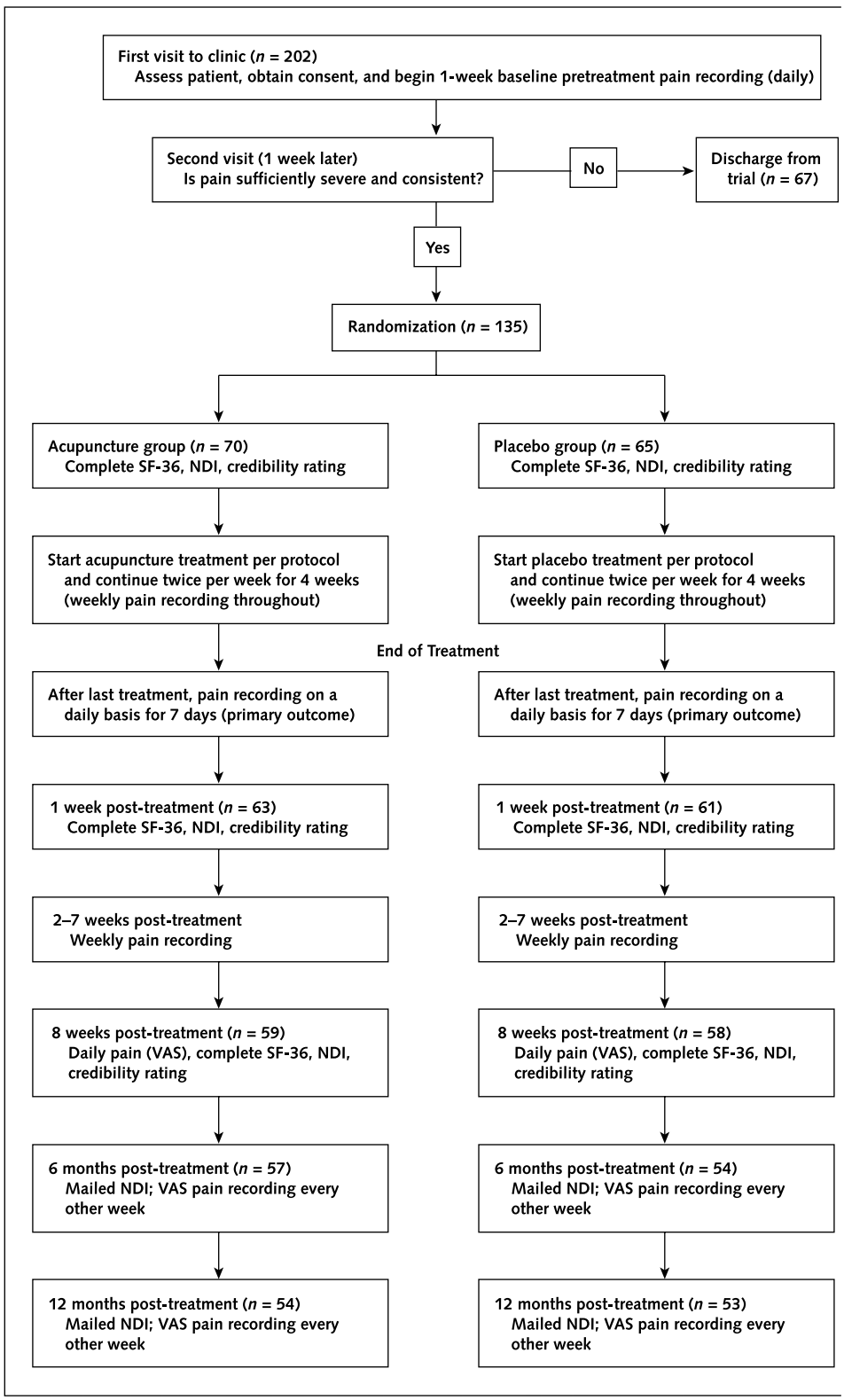
The Henry Smiths Charity and the Hospital Savings Association funded this project but had no role in the trial design or in dissemination of the results.

RESULTS

Patients

Between April 1999 and May 2001, 458 patients were considered for enrollment. Reasons for ineligibility were varied (for example, patients could not be contacted or

Figure 1. **Study flow with outcome assessments.**



NDI = Neck Disability Index; SF-36 = Short Form-36; VAS = visual analogue scale.

Table 1. **Baseline Characteristics***

Characteristic	Acupuncture Group (n = 70)	Control Group (n = 65)
Mean age, *y*	53.9 ± 15.71	52.8 ± 15.6
Men/women, *n/n*	24/46	24/41
Duration of disease, *y*†	4.81 ± 7.03	7.71 ± 11.39
Analgesic tablets, *n*	2.35 ± 2.42	2.21 ± 2.14
Baseline pain score on VAS, *mm*	49.6 ± 12.34	54.1 ± 14.6
NDI score	16.84 ± 6.34	17.18 ± 6.13
SF-36 PCS score	36.83 ± 7.91	36.33 ± 9.30
SF-36 MCS score	46.89 ± 10.38	48.30 ± 9.87
Credibility score	4.45 ± 1.1	4.55 ± 1.2
Spondylosis, *n (%)*	68 (97.1)	63 (96.9)
Other cause, *n (%)*‡	2 (2.9)	2 (3.1)
Fibromyalgia, *n (%)*	0	0
No previous experience of acupuncture, *n (%)*	66 (94.3)	58 (89.2)

* Values presented with a plus/minus sign are means ± SD. MCS = Mental Component Summary; NDI = Neck Disability Index; PCS = Physical Component Summary; SF-36 = Short Form-36; VAS = visual analogue scale.
† From onset of symptoms to date of trial.
‡ Prolonged idiopathic pain following trauma (whiplash).

were no longer in pain). Two hundred two patients were assessed, and 135 were randomly assigned (**Figure 1**), 70 to the acupuncture group and 65 to the placebo group. The most common reason for nonrandomization was failure to achieve a baseline VAS score of at least 30 mm.

Of the 135 patients randomly assigned to a study group, 11 (8.2%) withdrew during treatment. Three were men, and 8 were women; 7 (10%) withdrew from the acupuncture group, and 4 (6%) withdrew from the placebo group. The most common reason for withdrawal was the development or progression of another illness that required excluded medication (3 patients in the acupuncture group and 2 in the placebo group). Ninety-two percent of randomly assigned patients (n = 124) completed treatment and the first post-treatment assessments (**Figure 1**). We handled missing data in the daily assessments and questionnaires by averaging the available and relevant data on either side of the missing time point. However, this happened only rarely because few data were missing.

Twisk (38) investigated the robustness of generalized estimating equations and random coefficient analyses for various forms of missing values. For continuous dependent variables, he found that these techniques performed equally well with up to 25% of data missing. Because the amount of missing data in our study was much less than this, we did not use imputed data in the analyses.

Baseline Data

Patients were well balanced between groups at baseline (**Table 1**). The average daily VAS scores for pain were slightly lower in the acupuncture group (49.6 mm [95% CI, 47 mm to 53 mm]) than in the placebo group (54.1 mm [CI, 51 mm to 58 mm]). Duration of illness (defined as time elapsed between first onset of symptoms and entry into the trial) was slightly longer in the placebo group than in the acupuncture group. Differences in these baseline assessments were considered in the analyses.

Analysis of Outcome Variables

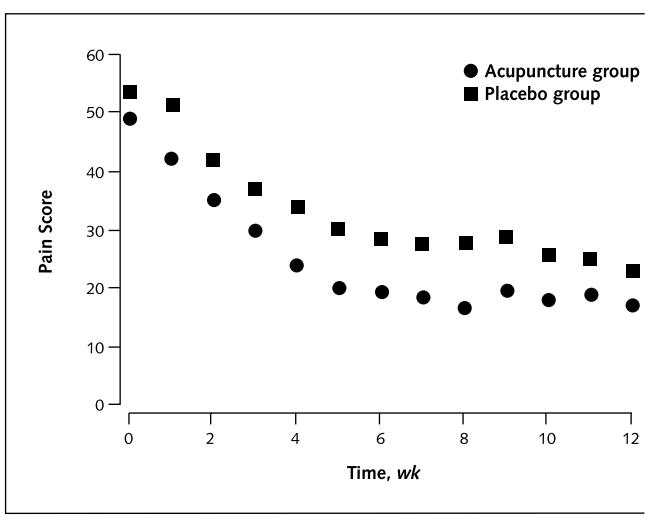
The pain scores for both groups decreased substantially (and statistically significantly) from baseline (**Table 2**, **Figure 2**). The initial model, fitted to the 5 weeks of pain scores, included possible confounders and interactions. Any insignificant interactions were removed from the model sequentially until the model shown in **Table 3** was produced. The patients' age and duration of illness at the start of the study were not predictors of subsequent pain. However, use of analgesic medications at baseline was associated with lower pain scores and higher analgesic use throughout the study was related to higher pain scores. Regardless of intervention group, patients with more pain at baseline reported more mean pain throughout the treatment period. The acupuncture group improved 6.3 mm (CI, 1.4 mm to 11.3 mm) more than the placebo group, a statistically significant difference (P = 0.01). The mean scores for both intervention groups decreased by about 4 mm for each week of treatment; this persisted into week 5. There was also a statistically significant hospital effect: The patients at Salisbury District Hospital improved an average of 7 mm more than patients at Southampton General Hospital, regardless of treatment allocation.

The pain scores for all 12 weeks of assessments were analyzed by using a similar longitudinal model (**Table 3**). Outcomes were similar at 5 and 12 weeks after randomization, and the treatment effect by group was almost the same (6 mm [CI, 0.9 to 10.9 mm]; P = 0.021). Women improved approximately 4 mm more than men.

Analysis of Secondary Outcome Variables

Overall, there were no statistically significant between-group differences for any of the secondary outcomes.

Figure 2. **Mean pain scores on the visual analogue scale from baseline to week 12.**



Treatment took place from week 1 to week 4.

Table 2. **Mean Pain and Acetaminophen Intake for Each Treatment Group at Major Time Points during the Trial***

Week†	Acupuncture Group				Placebo Group			
	Patients, *n*	Pain Score on VAS	Mean Daily Acetaminophen Tablets, *n*	Patients Using Acetaminophen, *n (%)*	Patients, *n*	Pain Score on VAS	Mean Daily Acetaminophen Tablets, *n*	Patients Using Acetaminophen, *n (%)*
Baseline	70	49.60 ± 12.35	2.9	56 (80)	65	54.11 ± 14.62	2.9	49 (75.4)
1	68	42.37 ± 17.56	2.1	45 (64.2)	64	51.59 ± 18.85	2.1	40 (61.6)
2	67	35.55 ± 19.18	1.9	37 (52.8)	62	42.26 ± 19.42	1.7	34 (52.3)
3	65	30.18 ± 22.80	2.0	31 (44.3)	62	37.39 ± 18.44	1.7	31 (47.6)
4	65	24.34 ± 21.63	2.1	28 (40.0)	61	34.38 ± 22.33	1.9	30 (46.1)
5	63	20.39 ± 20.26	2.6	34 (48.6)	61	30.69 ± 22.00	2.7	34 (52.3)
6	59	19.66 ± 18.36	1.6	25 (35.7)	59	28.83 ± 22.27	1.8	26 (40.0)
7	59	18.80 ± 19.11	1.8	23 (32.9)	59	27.98 ± 23.60	1.8	27 (41.6)
8	59	16.71 ± 17.73	1.9	21 (30.0)	59	28.20 ± 23.12	1.8	24 (37.0)
9	59	19.98 ± 22.43	1.6	27 (38.6)	59	29.03 ± 23.13	1.7	25 (38.5)
10	59	18.34 ± 20.81	2.0	19 (27.1)	59	26.10 ± 21.57	1.8	23 (35.4)
11	59	19.08 ± 21.67	1.6	23 (32.8)	59	25.27 ± 22.49	1.7	25 (38.5)
12	59	17.29 ± 18.96	2.4	27 (38.6)	58	23.19 ± 20.88	3.0	26 (40.0)
30	57	19.21 ± 24.22	–	–	54	21.02 ± 24.38	–	–
56	54	20.91 ± 25.69	–	–	53	24.36 ± 26.68	–	–

* Values presented with a plus/minus sign are means ± SD. VAS = visual analogue scale.
† Weeks 1 through 4 were treatment weeks.

Table 4 shows the means of NDI scores for the 2 study groups at baseline and at all post-treatment time points. The mean NDI score improved similarly from about 17 at baseline in both groups to approximately 9 and 11 in the acupuncture and placebo groups, respectively, at 12 months post-treatment. The difference between study groups was not statistically significant before or after multivariable adjustment.

The SF-36 was completed at baseline and at 5 and 12 weeks after randomization. The mean Physical Component Summary scores and Mental Component Summary scores are given in **Table 4**. Larger values indicate improvement, and for both components, scores were higher at 5 and 12 weeks among those who had had higher scores at baseline ($P < 0.001$). The Physical Component Summary score improved during the 12 weeks of the study from about 36 points to about 43 points (42.5 points and 43.8 points for the acupuncture and placebo groups, respectively), but there was no evidence of any treatment difference or any relationship with other covariates.

The Mental Component Summary score also improved during the first 5 weeks of treatment, to 61 points for both groups, but then deteriorated to pretreatment levels at 12 weeks. Multivariate analysis of the Mental Component Summary score showed that regardless of treatment group, women improved by 2.6 points (CI, 0.17 to 5.05 points) more at 5 weeks compared with men.

Credibility Ratings, Acetaminophen Use, and Adverse Reactions

No significant between-group differences were seen for any of the answers to the 4 questions used to assess credibility ($P > 0.2$ for all questions). This suggests that the 2 interventions had similar credibility before and after the treatment period. From the start of treatment, the number of patients taking acetaminophen decreased for both groups (**Table 2**), as did the mean number of tablets taken by each patient. Eight patients in the acupuncture group and 8 in the placebo group reported adverse reactions to the intervention (**Table 5**), including 4 patients who withdrew from treatment. There were no severe or serious adverse reactions.

Table 3. **Longitudinal Model Fitted to Pain Scores for Weeks 1 to 5 and Weeks 1 to 12**

Variable	Model Coefficient (95% CI)		*P* Value	
	Weeks 1 to 5	Weeks 1 to 12	Weeks 1 to 5	Weeks 1 to 12
Baseline pain	0.500 (0.300 to 0.701)	0.408 (0.201 to 0.615)	<0.001	<0.001
Baseline analgesic use	−1.316 (−2.946 to 0.315)	−1.657 (−3.180 to −0.134)	0.114	0.033
Analgesic use	0.622 (0.366 to 0.879)	0.723 (0.467 to 0.979)	<0.001	<0.001
Duration	0.075 (−0.141 to 0.291)	0.100 (−0.141 to 0.341)	>0.2	>0.2
Sex	−3.477 (−8.486 to 1.532)	−4.163 (−9.246 to 0.921)	0.174	0.109
Age	−0.073 (−0.231 to 0.084)	−0.030 (−0.183 to 0.124)	>0.2	>0.2
Hospital	−6.662 (−3.164 to −0.161)	−6.777 (−13.192 to −0.362)	0.045	0.038
Study week	−4.280 (−5.270 to −3.290)	−1.832 (−2.167 to −1.496)	<0.001	<0.001
Study group	6.335 (1.366 to 11.303)	5.906 (0.877 to 10.936)	0.012	0.021
Constant	32.853 (13.297 to 52.407)	31.374 (12.340 to 50.408)	0.001	<0.001

Table 4. **Scores on the Neck Disability Index and the Short Form-36 Physical Component Summary and Mental Component Summary at Baseline, 5 Weeks, 12 Weeks, 6 Months, and 12 Months***

Time Point	Acupuncture Group				Placebo Group			
	Patients, *n*	NDI Score	SF-36 PCS Score	SF-36 MCS Score	Patients, *n*	NDI Score	SF-36 PCS Score	SF-36 MCS Score
Baseline	70	16.84 ± 6.34	36.83 ± 7.91	46.89 ± 10.38	65	17.18 ± 6.13	36.33 ± 9.30	48.30 ± 9.87
1 wk post-treatment	64	11.78 ± 6.59	41.19 ± 7.89	61.07 ± 8.77	61	12.34 ± 7.35	40.58 ± 9.79	61.56 ± 9.05
8 wk post-treatment	59	10.98 ± 6.27	42.49 ± 9.75	52.49 ± 8.59	59	12.68 ± 7.79	43.75 ± 10.04	50.33 ± 10.13
6 mo post-treatment	56	9.91 ± 6.96	–	–	53	10.64 ± 8.3	–	–
12 mo post-treatment	53	8.89 ± 6.57	–	–	53	10.72 ± 9.11	–	–

* Values presented with a plus/minus sign are means ± SD. MCS = Mental Component Summary; NDI = Neck Disability Index; PCS = Physical Component Summary; SF-36 = Short Form-36.

DISCUSSION

Compared with baseline, acupuncture reduced pain by a mean of 58.9% at 1 week after treatment and by a mean of 65.1% at 8 weeks after treatment. This improvement was maintained in the longer term, with a decrease from baseline of 61.3% and 57.8% at 6 and 12 months after treatment, respectively. The longitudinal analysis of our primary outcome showed a significant difference between study groups at both 5 weeks ($P = 0.011$) and 12 weeks ($P = 0.005$) after randomization. The magnitude of the difference in pain between groups was 6 mm on the VAS, representing a between-group difference of 12% (CI, 3% to 21%) from baseline. This falls outside our definition of a clinically effective difference. Clinicians and researchers have varied, diverse, and somewhat arbitrary opinions about what constitutes a clinically significant effect in treatment of chronic pain. We based our power calculation on work by Lewith and Machin (17) and on pilot work by Petrie and Hazleman (21) and based our definition of a significant clinical outcome (30% between-group difference) on work by Melzack (39).

Double-blind trials of acupuncture are difficult, if not impossible, to construct. Therefore, if acupuncture is to be compared with placebo, patients' equipoise must be evaluated both before and after treatment assessment as a surrogate for blinding. In this study, both acupuncture and placebo were perceived as similarly credible, thus increasing the study's rigor and our confidence in our results.

Pain in both groups decreased by similar and statistically significant amounts, which suggests that the major improvement was not due to the specific acupuncture process. Furthermore, all of the secondary outcomes indicated substantial improvement over time with no differences between study groups other than pain. It should be noted, however, that some practitioners and experts in the field of acupuncture suggest that even light palpation of the acupuncture point (such as that applied when locating it to adhere an electrode) might constitute stimulation. If this is the case, our control might have been active, a possibility that, if true, would obviously affect our results. This would also have serious ramifications for acupuncture research in general because it would be impossible to keep patients from inadvertently stimulating any point themselves, for example, by rubbing an area of skin. Our trial estimates the

Table 5. **Adverse Reactions***

Patients	Age, *y*	Sex	Reaction
Acupuncture group			
1	48†	M	Increase in symptoms after 1 treatment
2	59	F	Faintness after the first treatment
3	62	F	Slight swelling of hand
4	27	F	Bruise at the site of LI 4
5	59	F	Faintness after the first treatment
6	49	F	Mild headache after each treatment
7	40†	F	Euphoria and enhanced or clearer vision post-treatment
8	71†	F	Dizziness after the first treatment
Placebo group			
1	37†	M	Discomfort during treatment; alleviated by decreasing intensity control
2	70	M	Mild headache after each treatment
3	69	F	Tiredness immediately after treatment on several occasions
4	61	F	Faintness after the 6th treatment
5	80	F	Nausea on several occasions after treatment
6	67	F	Tingling in the thumb during treatment
7	61	F	Dizziness after the first treatment
8	35	F	Uncomfortable cold feeling from the electrodes

* F = female; LI = large intestine; M = male.
† Withdrew from treatment.

specific and larger nonspecific effects of acupuncture treatment. It is vital to understand, particularly in view of observations by Beecher (40) and Kaptchuk and coworkers (41), whether this substantial nonspecific effect is unique to acupuncture or also applies to other physical treatments.

The vast majority of patients enrolled in our trial had experienced chronic pain from a degenerative condition for approximately 6 years. It is unlikely that the magnitude and tenacity of observed pain relief could be explained solely by regression to the mean (although this cannot be ruled out) or by the natural history of the illness. Pain did not decrease because of the confounding effects of increased analgesic use; the number of patients taking acetaminophen decreased for both groups and was substantially lower than at baseline. This study used longer-term follow-up than previous trials of acupuncture for neck pain. Its generalizability was further strengthened because we used a pragmatic treatment regimen, with an intervention based on the practice of western acupuncture in U.K. hospitals. This suggests that our results are relevant to many practitioners who use acupuncture.

Our trial raises several other interesting questions. Women tended to respond better than men on both the VAS and the SF-36. This requires further research to establish whether the sexes actually respond differently to acupuncture or whether the difference we observed was practitioner dependent. Generalizability of our results is impaired because only 1 practitioner provided treatment. More information and greater generalizability would be gained by using several practitioners in the same study. The Mental Component Summary score further highlights the importance of contact with a practitioner: Treatment cessation coincided with a dramatic score decrease. That patients from Salisbury District Hospital tended to achieve better results than patients from Southampton General Hospital might reflect a slight difference in hospital population, although we cannot suggest any obvious explanation other than chance. We do not know whether a Traditional Chinese Medicine–based approach might have yielded a different clinical outcome; therefore, we cannot comment on the "best" form of acupuncture.

Reviews of acupuncture unanimously call for more rigorous and methodologically sound studies. This paper reports such a trial, and we are confident that our results are neither spurious nor subject to major confounding factors. At present, there are few good trials against which ours can be compared. Perhaps the most comparable in terms of methodology are 2 early and very underpowered studies (42, 43) that yielded positive and negative results, respectively, in terms of efficacy.

Our well-powered trial clearly shows, through both the primary and the secondary outcomes, that patients in both study groups experienced significant, substantial, long-lasting improvement from baseline. For reducing neck pain, western acupuncture was statistically significantly more effective than placebo at both 5 and 12 weeks after randomization. This effect was not clinically significant in terms of our definition, however, even when taking into account the upper limits of the confidence intervals. The specific effect of acupuncture was relatively small in relation to its much larger nonspecific effect, a finding that has 2 main implications for clinicians and researchers. First, because acupuncture was clearly very effective at reducing pain, with patients experiencing large decreases over a prolonged period, its clinical use is recommended. Second, our study implies that most of the improvement gained from acupuncture was due not to the acupuncture process itself but predominantly to its powerful nonspecific effects. Further investigation of these nonspecific effects may lead to substantial enhancement of many clinical interventions.

From University of Southampton, Southampton, United Kingdom.

Acknowledgments: The authors thank Professor C. Cooper (Medical Research Council) and Dr. R. Ellis (Southampton General Hospital) for support with this project.

Grant Support: The study protocol was developed in 1997 and was funded by the Henry Smiths Charity and the Hospital Savings Association in 1998. Recruitment began in 1999 and was completed in 2001. Dr. Lewith's post was supported by the Laing Foundation.

Potential Financial Conflicts of Interest: None disclosed.

Requests for Single Reprints: Peter White, PhD, BSc, MCSP, Complementary Medicine Research Unit, Mail Primary Medical Care, University of Southampton, Aldermoor Health Centre, Aldermoor Close, Southampton SO16 5ST, United Kingdom; e-mail, pjw1@soton.ac.uk.

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

References

1. **Bogduk N.** The clinical anatomy of the cervical dorsal rami. Spine. 1982;7:319-30. [PMID: 7135065]

2. **Bogduk N, Lord SM.** Cervical spine disorders. Curr Opin Rheumatol. 1998;10:110-5. [PMID: 9567204]

3. **Sari-Kouzel H, Cooper R.** Managing pain from cervical spondylosis. Practitioner. 1999;243:334-8. [PMID: 10492978]

4. **White AR, Ernst E.** A systematic review of randomized controlled trials of acupuncture for neck pain. Rheumatology (Oxford). 1999;38:143-7. [PMID: 10342627]

5. **Berman BM, Swyers JP.** Establishing a research agenda for investigating alternative medical interventions for chronic pain. Prim Care. 1997;24:743-58. [PMID: 9386254]

6. **McCormack BM, Weinstein PR.** Cervical spondylosis. An update. West J Med. 1996;165:43-51. [PMID: 8855684]

7. **Thomas M, Eriksson SV, Lundeberg T.** A comparative study of diazepam and acupuncture in patients with osteoarthritis pain: a placebo controlled study. Am J Chin Med. 1991;19:95-100. [PMID: 1816730]

8. **Kjellman GV, Skargren EI, Oberg BE.** A critical analysis of randomised clinical trials on neck pain and treatment efficacy. A review of the literature. Scand J Rehabil Med. 1999;31:139-52. [PMID: 10458312]

9. **Laitinen J.** Treatment of cervical syndrome by acupuncture. Scand J Rehab Med. 1975;7:114-7.

10. **Aker PD, Gross AR, Goldsmith CH, Peloso P.** Conservative management of mechanical neck pain: systematic overview and meta-analysis. BMJ. 1996;313: 1291-6. [PMID: 8942688]

11. **Eisenberg DM, Davis RB, Ettner SL, Appel S, Wilkey S, Van Rompay M, et al.** Trends in alternative medicine use in the United States, 1990-1997: results of a follow-up national survey. JAMA. 1998;280:1569-75. [PMID: 9820257]

12. **Thomas KJ, Nicholl JP, Coleman P.** Use and expenditure on complementary medicine in England: a population based survey. Complement Ther Med. 2001;9:2-11. [PMID: 11264963]

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

14. **Barlas P.** In response to "Acupuncture for the relief of pain: a brief review" [Letter]. Physical Therapy Reviews. 1997;2:239-40.

15. **Paramore L.** Use of alternative therapies: estimates from the Robert Wood Johnson Foundation national access to care survey. J Pain Symptom Manage. 1996;13:83-9.

16. **White P, Lewith G, Berman B, Birch S.** Reviews of acupuncture for chronic neck pain: pitfalls in conducting systematic reviews. Rheumatology (Oxford). 2002;41:1224-31. [PMID: 12421994]

17. **Lewith GT, Machin D.** On the evaluation of the clinical effects of acupuncture. Pain. 1983;16:111-27. [PMID: 6348651]

18. **Ellis N.** Acupuncture in Clinical Practice. London: Chapman & Hall; 1999.

19. **Liangyue D, Yijun G, Shuhui H, Xiaoping J, Yang L, Rufen W, et al.** Chinese Acupuncture and Moxibustion. Beijing: Foreign Languages Pr; 1990.

20. **Dincer F, Linde K.** Sham interventions in randomized clinical trials of acupuncture—a review. Complement Ther Med. 2003;11:235-42. [PMID: 15022656]

21. **Petrie J, Hazleman B.** Credibility of placebo transcutaneous nerve stimulation and acupuncture. Clin Exp Rheumatol. 1985;3:151-3. [PMID: 4017314]

22. **de la Torre CS.** The choice of control groups in invasive clinical trials such as acupuncture. Frontier Perspectives. 1993;3:33-7.

23. **Wood R, Lewith G.** The credibility of placebo controls in acupuncture studies. Complement Ther Med. 1998;6:79-82.

24. **Riddle DL, Stratford PW.** Use of generic versus region-specific functional status measures on patients with cervical spine disorders. Phys Ther. 1998;78: 951-63. [PMID: 9736893]

25. **Vernon H, Mior S.** The Neck Disability Index: a study of reliability and validity. J Manipulative Physiol Ther. 1991;14:409-15. [PMID: 1834753]

26. **Vernon H.** The Neck Disability Index: patient assessment and outcome monitoring in whiplash. Journal of Musculoskeletal Pain. 1996;4:95-104.

27. **Hains F, Waalen J, Mior S.** Psychometric properties of the neck disability index. J Manipulative Physiol Ther. 1998;21:75-80. [PMID: 9502061]

28. **Westaway MD, Stratford PW, Binkley JM.** The patient-specific functional scale: validation of its use in persons with neck dysfunction. J Orthop Sports Phys Ther. 1998;27:331-8. [PMID: 9580892]

29. **Beaton DE, Hogg-Johnson S, Bombardier C.** Evaluating changes in health status: reliability and responsiveness of five generic health status measures in workers with musculoskeletal disorders. J Clin Epidemiol. 1997;50:79-93. [PMID: 9048693]

30. **Andresen EM, Rothenberg BM, Panzer R, Katz P, McDermott MP.** Selecting a generic measure of health-related quality of life for use among older adults. A comparison of candidate instruments. Eval Health Prof. 1998;21:244-64. [PMID: 10183346]

31. **Hemingway H, Stafford M, Stansfeld S, Shipley M, Marmot M.** Is the SF-36 a valid measure of change in population health? Results from the Whitehall II Study. BMJ. 1997;315:1273-9. [PMID: 9390054]

32. **Shapiro ET, Richmond JC, Rockett SE, McGrath MM, Donaldson WR.** The use of a generic, patient-based health assessment (SF-36) for evaluation of patients with anterior cruciate ligament injuries. Am J Sports Med. 1996;24:196-200. [PMID: 8775120]

33. **Millard RW, Beattie PF, Jones RH.** A comprehensive review of questionnaires to evaluate chronic pain-related disability. Critical Reviews in Physical and Rehabilitation Medicine. 1997;9:35-52.

34. **Vincent C.** Credibility assessment in trials of acupuncture. Complementary Medicine Research. 1990;4:8-11.

35. **Vincent C, Lewith G.** Placebo controls for acupuncture studies. J R Soc Med. 1995;88:199-202. [PMID: 7745565]

36. **Zeger SL, Liang KY.** Longitudinal data analysis for discrete and continuous outcomes. Biometrics. 1986;42:121-30. [PMID: 3719049]

37. **Zeger SL, Liang KY, Albert PS.** Models for longitudinal data: a generalized estimating equation approach. Biometrics. 1988;44:1049-60. [PMID: 3233245]

38. **Twisk J.** Applied Longitudinal Data Analysis for Epidemiology—A Practical Guide. New York: Cambridge Univ Pr; 2003.

39. **Melzack R.** Prolonged relief of pain by brief, intense transcutaneous somatic stimulation. Pain. 1975;1:357-73. [PMID: 141644]

40. **Beecher H.** The powerful placebo. JAMA. 1955;159:1602-6.

41. **Kaptchuk TJ, Goldman P, Stone DA, Stason WB.** Do medical devices have enhanced placebo effects? J Clin Epidemiol. 2000;53:786-92. [PMID: 10942860]

42. **Petrie JP, Langley GB.** Acupuncture in the treatment of chronic cervical pain. A pilot study. Clin Exp Rheumatol. 1983;1:333-6. [PMID: 6335859]

43. **Petrie JP, Hazleman BL.** A controlled study of acupuncture in neck pain. Br J Rheumatol. 1986;25:271-5. [PMID: 3524727]

Appendix Table. **Point Selection**

Primary Local Neck Points	Secondary, Alternative Local Neck Points	Primary Distal Points	Secondary, Alternative Distal Points
GB 20, 21; GV 14	**SI 12, 13 14; BL 9, 10;** ST 11; SI 15, 16; BL 11, 41, 15, 17; GB 29; TE 16, 17; GV 15, 16, 17	**LI 4, SI 3, GB 34, TE 5**	LI 11; SI 8; TE 10, 36, 39, 40; BL 60; extra Luozhen

* Points in boldface are the most commonly used and recommended. BL = bladder; GB = gallbladder; GV = govenor vessel; LI = large intestine; SI = small intestine; ST = stomach; TE = triple energizer.

Current Author Addresses: Drs. White and Lewith: Complementary Medicine Research Unit, Mail Primary Medical Care, University of Southampton, Aldermoor Health Centre, Aldermoor Close, Southampton SO16 5ST, United Kingdom.
Dr. Prescott: School of Mathematics, University of Southampton, Highfield, Southampton SO17 1BJ, United Kingdom.
Dr. Conway: School of Health Professions, University of Southampton, Highfield, Southampton SO17 1BJ, United Kingdom.

Author Contributions: Conception and design: P. White, G. Lewith, J. Conway.
Analysis and interpretation of the data: P. White, G. Lewith, P. Prescott, J. Conway.
Drafting of the article: P. White, G. Lewith, P. Prescott.
Critical revision of the article for important intellectual content: P. White, G. Lewith.
Final approval of the article: P. White, G. Lewith.
Statistical expertise: G. Lewith, P. Prescott.
Obtaining of funding: G. Lewith.
Administrative, technical, or logistic support: G. Lewith, J. Conway.
Collection and assembly of data: P. Prescott.