Acupuncture for Chronic Low Back Pain in Long-Term Follow-Up: A Meta-Analysis of 13 Randomized Controlled Trials

Mai Xu,* Shi Yan,† Xu Yin,§ Xiuyang Li,‡ Shuguang Gao,* Rui Han,¶ Licheng Wei,* Wei Luo* and Guanghua Lei*

*Department of Orthopaedics, Xiangya Hospital, Central South University, Changsha, Hunan 410008, China
†The Institute of Documents on Traditional Chinese Medicine
‡Department of Acupuncture and Moxibustion, Shandong University of Traditional Chinese Medicine, Jinan, Shandong 250000, China
§Department of Pathology in Basic Medicine, Beijing University of Traditional Chinese Medicine, Chaoyang, Beijing 100029, China
¶Department of Orthopaedics, Qingdao Haici Hospital, Qingdao, Shandong 266000, China

Abstract: Chronic low back pain is one of the most common reasons that people seek medical treatment, and the consequent disability creates a great financial burden on individuals and society. The etiology of chronic low back pain is not clear, which means it is often refractory to treatment. Acupuncture has been reported to be effective in providing symptomatic relief of chronic low back pain. However, it is not known whether the effects of acupuncture are due to the needling itself or nonspecific effects arising from the manipulation. To determine the effectiveness of acupuncture therapy, a meta-analysis was performed to compare acupuncture with sham acupuncture and other treatments. Overall, 2678 patients were identified from thirteen randomized controlled trials. The meta-analysis was performed by a random model (Cohen’s test), using the I-square test for heterogeneity and Begg’s test to assess for publication bias. Clinical outcomes were evaluated by pain intensity, disability, spinal flexion, and quality of life. Compared with no treatment, acupuncture achieved better outcomes in terms of pain relief, disability recovery and better quality of life, but these effects were not observed when compared to sham acupuncture. Acupuncture achieved better outcomes when compared with other treatments. No publication bias was detected.

Correspondence to: Dr. Guanghua Lei, Department of Orthopaedics, Xiangya Hospital, Central South University, No. 87 XiangYa Road, Changsha, Hunan 410008, China. Tel: (+86) 731-8432-7326, Fax: (+86) 731-8432-7354, E-mail: guanghualei9640@gmail.com
Acupuncture is an effective treatment for chronic low back pain, but this effect is likely to be produced by the nonspecific effects of manipulation.

Keywords: Meta-Analysis; Chronic Low Back Pain; Efficacy; Disability; Acupuncture.

Introduction

Low back pain, which carries a 70–85% lifetime risk in the general population (Andersson, 1999), is been one of the most common reasons for people to seek medical treatment (Hart et al., 1995; Deyo et al., 2006). It also creates a great financial burden as a result of the consequent disability and the utilization of health services (Frymoyer and Cats-Baril, 1991; Frymoyer, 1992; Waddell, 1996; 2010, Apeldoorn et al., 2010; Wieser et al., 2010; Guidelines for lower back pain treatment ignored, 2010). Chronic low back pain is a poorly defined condition as the underlying pathological causes are often unclear (Mayer et al., 2010). Some physiotherapists and researchers have argued that the chronic disability in low back pain should be primarily considered as a psychosocial dysfunction (Cai et al., 2007; Astfalck et al., 2010). Conventional medication is often ineffective in its treatment, whereas alternative therapies have sometimes been shown to achieve excellent outcomes in terms of both pain control and rehabilitation (Kumar et al., 2009; Witt et al., 2009; Sherman et al., 2010).

Acupuncture is an ancient therapy that originates from traditional Chinese philosophy, where the disease is considered to occur as a consequence of the imbalance between the Yin and Yang forces in the body. Inserting needles into the specific acupoints on the meridian lines can restore the balance of the body. Many studies have confirmed the effect of acupuncture on pain (Kaptchuk, 2002; Thicke et al., 2011). However, most of these findings cannot be replicated due to insufficient sample sizes or limitations in the methodology, and the utility of acupuncture in modern medicine remains controversial. Meta-analyses have therefore been applied to explore the effects of pain relief. Various systematic reviews regarding the use of acupuncture for back pain have been published (Ernst and White, 1998; Van Tulder et al., 1999; Furlan et al., 2005a, 2005b; Manheimer et al., 2005; Yuan et al., 2008; Kong et al., 2009; Chou, 2010; Linde et al., 2010; Rubinstein et al., 2010; Furlan et al., 2012). Some have confirmed that acupuncture has a marginally greater effect in relieving pain compared to sham acupuncture and no additional treatment (Manheimer et al., 2005). Others showed that acupuncture could only improve pain relief and functional improvement immediately after the sessions and within a short-term follow-up period, compared to no treatment or sham therapy (Furlan et al., 2005b). Although the local and endocrine control mechanisms of acupuncture have been clearly explained (Luu and Boureau, 1989; Chu, 2002; Cai et al., 2009; Wang et al., 2010), its impact on chronic low back pain has inevitably been attributed to non-specific effects (Linde et al., 2010).

We raise the null hypothesis that acupuncture is equally as effective as blank treatments, sham acupuncture, or conventional care, as well as other alternative therapies, such as transcutaneous electrical nerve stimulation and massage. We attempted to confirm this
hypothesis with a meta-analysis of the literature, which demonstrated that acupuncture is
effective for chronic low back pain due to the non-specific effects that arise from the
manipulation.

Materials and Methods

Review Objective

This meta-analysis is reported according to updated methodological guidelines for sys-
tematic reviews published by the Cochrane back review group (2009) (Chou, 2010). The following comparisons were investigated: (1) acupuncture compared with blank
treatment in a relevant comparison (i.e., a comparison of acupuncture with another treat-
ment versus the same other treatment alone) or an irrelevant comparison (a comparison of
acupuncture versus no treatment) to test whether acupuncture is useful for the treatment of
chronic low back pain; (2) acupuncture compared with sham acupuncture to test the
non-specific effects that arise from manipulation; and (3) acupuncture compared with other
conventional or alternative treatments to test whether acupuncture is superior to these other
treatments.

Search Strategy

Two reviewers independently performed a computer-aided search on Pubmed/Medline,
AMED, EMBASE, Cochrane Central Register of Controlled Trials (CENTRAL), ISI
Proceedings for Conference Abstracts, International Standard Randomized Controlled Trial
Number (ISRCTN) Register and the meta-Register for Randomized Controlled Trials
(mRCT), China National Knowledge Infrastructure (CNKI) and the Wan Fang database up
to 25 January 2012. A search of specific databases was also performed, included the
Complementary and Alternative Medicine Specialist Library (from the National Library of
Health, UK), the Physiotherapy Evidence Database (PEDro), and the Cumulative Index of
Nursing and Allied Health (CINAHL).

In order to find as many studies as possible, we used a loose search strategy. Search
terms included “acupuncture”, “low back pain”, “needle”, “back pain” with the limits set
for randomized controlled trials (RCTs). The terms were searched both as free terms and
Medical Subject Headings (MeSH) terms and connected by the Boolean operators “AND”
and “OR”. Titles and abstracts of the potentially involved articles were reviewed first, to
decide whether the studies contained any relevant information. If the trial was deemed to be
eligible by either author, the full text would be obtained for further review. Keywords of
the references of all relevant publications were traced, including systematic reviews, meta-
analyses and identified RCTs, in order to retrieve citations omitted by the electronic search.
We then sent a copy of the selected studies to all authors of the studies involved to ask if
they were acknowledged in any other published trials. The search was repeated near the
completion of the data analysis (15 February 2012) to ensure an integrated investigation.
Figure 1 details the trial search flow of our meta-analysis.
Inclusion Criteria

We included therapeutic RCTs, and pain intensity was established as the primary outcome, while the secondary outcomes were the indexes related to the clinical outcome (e.g., disability, quality of life and spinal flexion). After the literature review, we found that the definition of the term “chronic” in terms of low back pain was often confusing. Thus, we primarily included articles in the meta-analysis that defined “chronic patients” as those who had suffered from pain for at least the last three weeks (Rubinstein et al., 2010), and then secondarily those that had suffered pain for 12 weeks (Chou, 2010). “Acupuncture” was defined as a process wherein needles are penetrated into the skin without an injection, accompanying a definite feeling of “De Qi” (an energetic feedback that the practitioner feels when reaching the correct placement of an acupuncture needle). “Sham” acupuncture was defined as either puncturing a location near the acupoint with tingling only but not “De Qi”, or simulated acupuncture technique using a toothpick or other needle-like object in the needle guidetube (Cherkin et al., 2009). This has been found to be effective in acupuncture-naïve patients with back pain (Sherman et al., 2002). Long-term follow-up refers to outcomes that were measured between four weeks and one year (Chou, 2010).

When a single same trial was described in more than one publication, we only included the latest report. We excluded studies with non-clinical outcome measures (e.g., a study on the mechanism), comparisons between the various acupuncture methods (e.g., moxibustion, electroacupuncture, warm needles, with or without Thermal Design Power), three-dimensional finite element analysis, animal studies, case reports, comparative studies without randomization, expert opinions, editorials, letters, practice guidelines and reviews.
Risk of Bias and Studies Quality Evaluation

The risk of bias was appraised in accordance with the Cochrane Back Review Group Criteria List for Methodologic Quality Assessment of Randomized, Controlled Trials (van Tulder et al., 2003). The report quality of the acupuncture intervention assessment was made based upon revised Standards for Reporting Interventions in Clinical Trials of Acupuncture (STRICTA). Disagreements or inconsistencies were resolved through discussion in cooperation with a third author (XY L).

Data Extraction

Two authors (SY and XY) independently extracted data from each included study in duplicate with the use of a standardized form. Data for general information (author, publication year), mean follow-up duration, minimal or mean pain duration, mean patient age, sex distribution and the number of patients were listed. The sample size was halved if the acupuncture group was compared with two treatments. Again, disagreements were settled by discussion between the review authors and if necessary with help from a third author.

Statistical Methods

We entered eligible trials into Stata/SE 11.0 software. The I-square score was used to quantify heterogeneity between trials, and > 50% was considered as significantly heterogeneous. The trial results were factitiously divided into primary outcome measures (pain intensity) and secondary outcome measures (pain disability, spine flexion and quality of life). When an interval data was the only available data to describe the discrete trend, we used approximations to estimate standard deviation (SD) as 95% confidence interval (CI) divided by “4” or (75–25% CI) divided by “2”. We employed a random model to meta-analyze the end points when a high percentage was obtained for I^2 (> 50%). The standard mean difference (SMD) was obtained as the effect size by Cohen’s test and associated 95% CIs. The effect size was defined as 0.20 for small, 0.50 for medium, and 0.80 for large effects (Cohen, 1992). Stratified analysis by treatment was used to determine whether the outcomes were stable. For pain intensity, pain disability and spinal flexion, negative effect sizes indicated a beneficial effect of acupuncture, while quality of life was antithetical. Begg’s rank test was used to assess publication bias. Consistency in direction was defined as when 75% or more of the studies showed either a benefit or no benefit (Chou, 2010).

Results

Study Characteristics of Eligible Studies and Quality Assessment

We considered 132 potentially relevant references. Sixty-six studies were excluded after the titles and abstracts were reviewed. Twenty-three studies were excluded because of
duplication and 34 more were excluded as a result of an evaluation index of no interest (e.g., adverse events and psychological distress). A small amount of the index could not be merged (Coan et al., 1980) or had an inappropriate study design, such as analgesic intake and activity improvement (Mendelson et al., 1983; Giles and Muller, 2003). An additional manual search of reference lists generated nine further citations that were not obtained from the electronic search, of which five were excluded because of the language (Korean). After the application of our inclusion and exclusion criteria, 2678 patients were identified from 13 articles (Table 1) (Grant et al., 1999; Leibing et al., 2002; Molsberger et al., 2002; Kerr et al., 2003; Meng et al., 2003; Yeung et al., 2003; Brinkhaus et al., 2006; Thomas et al., 2006; Haake et al., 2007; Szczurko et al., 2007; Cherkin et al., 2009; Itoh et al., 2009; Zaringhalam et al., 2010). Of the thirteen trials, one was from Northern Ireland, one from Hong Kong, one from Japan, one from Canada, one from Iran, two from the United Kingdom, two from the USA, and four from Germany.

The demographic data are shown in Table 1. The acupuncture invention assessment and risk of bias assessment are listed in Tables 2 and 3, respectively. Most of the involved studies had flaws in the reporting style for acupuncture, the depth of needle insertion, needle retention times, the setting and context of treatment, and the background of the practitioner. The main biases that affected the results were performance bias and detection bias. Begg’s rank test showed that there was no publication bias in terms of pain intensity and pain disability, with \( p \) values (continuity corrected) of 0.544 and 0.537, respectively. Figures 2A and 2B show the Begg’s funnel plots that visualize the tests for publication bias regarding pain intensity and disability as a result of pain.

Table 1. Details of Trials of Acupuncture versus Sham Acupuncture, Blank and Other Treatments for Chronic Low Back Pain

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Mean of Follow-Up</th>
<th>Duration of Pain</th>
<th>No. of Patients</th>
<th>Mean Age</th>
<th>Sex Ratio (M:F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leibing E</td>
<td>2002</td>
<td>9m</td>
<td>NA*</td>
<td>131</td>
<td>48.1y</td>
<td>55:76</td>
</tr>
<tr>
<td>Molsberger AF</td>
<td>2002</td>
<td>3m</td>
<td>Mean 9.9 yr</td>
<td>186</td>
<td>50y</td>
<td>97:89</td>
</tr>
<tr>
<td>Meng CF</td>
<td>2003</td>
<td>4w</td>
<td>( \geq 12 ) wk</td>
<td>47</td>
<td>71y</td>
<td>18:29</td>
</tr>
<tr>
<td>Kerr DP</td>
<td>2003</td>
<td>6m</td>
<td>NA*</td>
<td>60</td>
<td>41y</td>
<td>28:32</td>
</tr>
<tr>
<td>Brinkhaus B</td>
<td>2006</td>
<td>52w</td>
<td>14.7 yr</td>
<td>298</td>
<td>58.8y</td>
<td>96:202</td>
</tr>
<tr>
<td>Thomas KJ</td>
<td>2006</td>
<td>24m</td>
<td>NA*</td>
<td>239</td>
<td>NA*</td>
<td>94:145</td>
</tr>
<tr>
<td>Cherkin DC</td>
<td>2009</td>
<td>52w</td>
<td>3–12 mo</td>
<td>638</td>
<td>47y</td>
<td>242:396</td>
</tr>
<tr>
<td>Yeung CKN</td>
<td>2003</td>
<td>3m</td>
<td>NA*</td>
<td>52</td>
<td>53y</td>
<td>9:43</td>
</tr>
<tr>
<td>Grant DJ</td>
<td>1999</td>
<td>3m</td>
<td>( \geq 6 ) mo</td>
<td>60</td>
<td>74y</td>
<td>6:54</td>
</tr>
<tr>
<td>Kazunori I</td>
<td>2009</td>
<td>12w</td>
<td>( \geq 6 ) mo</td>
<td>26</td>
<td>( &gt; 65 ) y</td>
<td>9:17</td>
</tr>
<tr>
<td>Szczurko O</td>
<td>2007</td>
<td>12w</td>
<td>( \geq 6 ) weeks</td>
<td>75</td>
<td>46.6y</td>
<td>NA*</td>
</tr>
<tr>
<td>Haake M</td>
<td>2007</td>
<td>6m</td>
<td>8 years</td>
<td>1162</td>
<td>50y</td>
<td>470:692</td>
</tr>
<tr>
<td>Zaringhalam J</td>
<td>2010</td>
<td>10w</td>
<td>NA*</td>
<td>80</td>
<td>50–60y</td>
<td>all men</td>
</tr>
</tbody>
</table>

Note: *NA = data not available.
Table 2. Acupuncture Intervention Assessment According to the Revised Standards for Reporting Interventions in Clinical Trials of Acupuncture (STRICTA)

|               | 1a | 1b | 1c | 2a | 2b | 2c | 2d | 2e | 2f | 2g | 3a | 3b | 4a | 4b | 5 | 6a | 6b | Total Score |
|---------------|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|-------------|
| Leibing E     | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 17          |
| Molsberger AF  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 0  | 0  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 15          |
| Meng CF       | 0  | 1  | 1  | 1  | 1  | 0  | 1  | 1  | 1  | 1  | 1  | 0  | 1  | 1  | 1  | 1  | 1  | 13          |
| Kerr DP       | 0  | 1  | 0  | 1  | 0  | 1  | 0  | 1  | 1  | 1  | 1  | 1  | 0  | 1  | 1  | 1  | 1  | 11          |
| Brinkhaus B   | 0  | 0  | 1  | 1  | 1  | 0  | 1  | 0  | 0  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 12          |
| Thomas KJ     | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 1  | 1  | 0  | 0  | 1  | 1  | 1  | 5           |
| Cherkin DC    | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 17          |
| Yeung CKN     | 0  | 1  | 0  | 1  | 1  | 1  | 1  | 1  | 0  | 1  | 1  | 1  | 1  | 0  | 1  | 1  | 1  | 13          |
| Grant DJ      | 1  | 1  | 1  | 1  | 0  | 0  | 0  | 0  | 1  | 1  | 1  | 1  | 1  | 1  | 0  | 1  | 1  | 11          |
| Kazunori I    | 0  | 0  | 1  | 1  | 0  | 0  | 0  | 1  | 1  | 1  | 1  | 1  | 0  | 0  | 0  | 1  | 1  | 9           |
| Szczurko O    | 0  | 0  | 1  | 1  | 1  | 1  | 1  | 0  | 1  | 1  | 1  | 1  | 0  | 0  | 1  | 1  | 1  | 12          |
| Haake M       | 0  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 0  | 1  | 1  | 1  | 1  | 1  | 0  | 1  | 1  | 14          |
| Zarighalam J  | 0  | 1  | 0  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 0  | 0  | 1  | 1  | 1  | 13          |

1: ACUPUNCTURE RATIONALE
1a: Style of acupuncture (e.g., Traditional Chinese Medicine (TCM), Japanese, Korean, Western medical, Five Element, ear acupuncture, etc).
1b: Reasoning for treatment provided, based on historical context, literature sources and/or consensus methods, with references where appropriate.
1c: Extent to which treatment was varied.

2: DETAILS OF NEEDLING
2a: Number of needle insertions per subject per session (mean and range where relevant).
2b: Names (or location if no standard name) of points used (uni-/bi-lateral).
2c: Depth of insertion, based on a specified unit of measurement or on a particular tissue level.
2d: Response sought (e.g., De Qi or muscle twitch response).
2e: Needle stimulation (e.g., manual, electrical).
2f: Needle retention time.
2g: Needle type (diameter, length and manufacturer or material).

3: TREATMENT REGIMEN
3a: Number of treatment sessions.
3b: Frequency and duration of treatment sessions.

4: OTHER COMPONENTS OF TREATMENT
4a: Details of other interventions administered to the acupuncture group (e.g., moxibustion, cupping, herbs, exercises, lifestyle advice).
4b: Setting and context of treatment, including instructions to practitioners, and information and explanations to patients.

5: PRACTITIONER BACKGROUND
5: Description of participating acupuncturists (qualification or professional affiliation, years in acupuncture practice, other relevant experience).

6: CONTROL OR COMPARATOR INTERVENTIONS
6a: Rationale for the control or comparator in the context of the research question, with sources that justify the choice(s).
6b: Precise description of the control or comparator. If sham acupuncture or any other type of acupuncture-like control is used, provide details as for items 1–3 above.
Our primary method was to compare the outcome of acupuncture versus blank treatments, sham acupuncture and other treatments as reported in all 13 involved studies using meta-analysis.

**Pain intensity.** Figure 3 shows the results of pain intensity changes. Pain intensity was reported in 18 studies from 13 articles, in which eight compared acupuncture with sham acupuncture, five with a group blank and five with other treatments. Visual analog scales (VAS) were employed in nine articles; the studies by Cherkin et al. (2009) and Yeung et al. (2003) used the numerical rating scale, Haake et al. (2007) used the Von Korff Chronic Pain Grade Scale and Thomas et al. (2006) used the McGill present pain index, all of
which are very similar scoring systems to the VAS. The heterogeneity across the subgroup studies was high. The overall SMD was $-0.43$ (95% CI, $-0.64$ to $-0.21$); the blank treatment and other treatment subgroups were significantly different, with a SMD of $-0.64$ (95% CI, $-1.13$ to $-0.14$) and $-0.49$ (95% CI, $-0.90$ to $-0.09$), respectively. Although the effect size showed that acupuncture could relieve pain more effectively than sham treatment (SMD, $-0.26$; 95% CI, $-0.56$ to 0.05), this finding should not be considered statistically significant as the 95% CI crossed 0. The effect size was small when acupuncture was compared with other treatments and moderate when compared with the blank treatment. The subgroup effect sizes are shown in Fig. 4A.
Disability as a result of pain. Figure 5 shows the results regarding disability as a result of pain. Data regarding this parameter were reported in 12 studies from nine articles, in which four compared acupuncture with sham acupuncture, four compare acupuncture with the blank treatment group and four with other treatments. The Roland disability questionnaire was employed in the studies by Meng et al. (2003), Zaringhalam et al. (2010) and Cherkin et al. (2009), while the pain disability index was used by Leibing et al. (2002) and Brinkhaus et al. (2006), Kazunori et al. (2009) Zaringhalam et al. (2010) and Thomas et al. (2006) used the Oswestry pain disability index, and the Aberdeen low back pain scale was used by in Yeung et al. (2003). All the evaluation indexes were converted into a hundred percentage point system.

The heterogeneity across the subgroup studies was low, with the exception of the other treatment subgroup. The SMD of the overall cohort was $-0.43$ (95% CI, $-0.66$ to $-0.21$); Acupuncture in both the blank and other treatment subgroups gained better outcomes as the SMD was $-0.58$ (95% CI, $-0.82$ to $-0.34$) and $-0.71$ (95% CI, $-1.29$ to $-0.21$). The effect size of acupuncture was moderate compared with the other treatment and the blank treatment groups. However, compared with the sham treatment group, acupuncture was not statistically different. The subgroup effect sizes are shown in Fig. 4B.

![Figure 3. Meta-analysis for pain intensity. The diamond represents the summary treatment effect over all analyzed studies. A negative SMD represents that the acupuncture is superior to control intervention in relieving pain.](image-url)
Spinal flexion. Figure 6 shows the results regarding spinal flexion. Data were reported in five studies, of which three compared acupuncture with sham acupuncture, one with blank treatment and one with other treatments. The fingertip-to-floor distance is employed in three studies while the range of movement of spinal flexion was assessed in two studies. Heterogeneity was low in the sham treatment subgroup (I² = 0.0%) and high overall (I² = 77.4%). The other treatments subgroup was the only one that significantly favored acupuncture, with an SMD of −1.04 (95% CI, −1.56 to −0.51). The effect size of this finding was high. The subgroup effect sizes are shown in Fig. 4C.

Quality of life. Quality of life was measured by physical and mental functioning scoring systems. Relevant data was available in five studies, of which three compared acupuncture with sham acupuncture, one with blank treatment and one with other treatments. The evaluation indexes used were SF-12 and SF-36. Figure 7 shows the contribution of each study to the meta-analysis. Heterogeneity was low in the sham subgroup (I² = 17.0%), but the overall effect was significantly heterogeneous (I² = 85.1%). When acupuncture was compared with sham acupuncture, the SMD was significantly in favor of acupuncture (SMD, 0.22; 95% CI, 0.03 to 0.40) but the effect size was still smaller than the overall (SMD, 0.47; 95% CI, 0.15 to 0.78), blank treatment...
Figure 5. Meta-analysis for disability. The diamond represents a summary of the treatment effect across all analyzed studies. A negative SMD demonstrates that acupuncture is superior to the control intervention for recovering from disability.

Figure 6. Meta-analysis for spinal flexion. The diamond represents a summary of the treatment effect across all analyzed studies. The overall SMD is not statistically significant as the 95% confidence interval includes 0. This corresponds to a \( p \) value > 0.05.
Secondary Meta-Analysis

We secondarily meta-analyzed the outcomes of acupuncture versus blank, sham acupuncture and other treatments as reported in the studies that clearly defined chronic pain as pain which lasted longer than 12 weeks (Grant et al., 1999; Molsberger et al., 2002; Meng et al., 2003; Brinkhaus et al., 2006; Haake et al., 2007; Cherkin et al., 2009; Itoh et al., 2009). The results are listed in Figs 4E–4H. Compared with sham acupuncture, acupuncture showed no superior benefit in the treatment of chronic low back pain in all four evaluation indexes (pain intensity, disability, spinal flexion and quality of life). Furthermore, acupuncture only showed a better outcome in terms of disability and quality of life when compared with the blank and other treatments only in pain disability.

Consistency in Direction

The consistency in direction of pain intensity and quality of life changes was poor, as only 55.5% (10/18) and 60% (3/5) of the studies showed a trend that was consistent with subgroup effects, respectively. However, the consistency was good for disability and spinal flexion, as 75% (9/12) and 100% (5/5) of the studies showed a trend that was consistent with subgroup effects, respectively.
Consistent results were obtained in 72% (8/11 and 13/18) of the studies in the blank and sham acupuncture groups, while consistency was poor in the other treatment group (6/11, 55%).

Discussion

Numerous systemic reviews and meta-analyses over the last two decades have investigated the effectiveness of acupuncture for the treatment of chronic low back pain. When acupuncture has been compared with no additional treatment or blank treatments, the results have consistently shown that acupuncture is significantly more effective (Ernst and White, 1998; Furlan et al., 2005a, 2005b; 2012; Manheimer et al., 2005; Yuan et al., 2008; Linde et al., 2010; Rubinstein et al., 2010). However, the results are relatively inconsistent when acupuncture is compared with sham acupuncture (Ernst and White, 1998; Van Tulder et al., 1999; Furlan et al., 2005a, 2005b; 2012; Manheimer et al., 2005; Yuan et al., 2008; Kong et al., 2009; Linde et al., 2010; Rubinstein et al., 2010). Some studies found there is moderate evidence that acupuncture is more effective than sham acupuncture (Van Tulder et al., 1999; Furlan et al., 2005a, 2005b; Manheimer et al., 2005; Linde et al., 2010; Rubinstein et al., 2010), while other studies have published strong evidence of no significant difference between acupuncture and sham acupuncture (Ernst and White, 1998; Yuan et al., 2008; Kong et al., 2009; Furlan et al., 2012). Given the clinical heterogeneity of other treatments for chronic low back pain, it is not surprising that a consistent conclusion could not be made when comparing acupuncture with other treatments: some are positive (Furlan et al., 2005a; Yuan et al., 2008; Rubinstein et al., 2010; Furlan et al., 2012) and some are negative (Van Tulder et al., 1999). Drawing a generalized conclusion is complicated by the fact that different systemic reviews and meta-analyses have used different inclusion and exclusion criteria, different definitions of “chronic pain” and “acupuncture”, and evaluated different types of outcomes.

Acupuncture is Effective in Relieving Chronic Low Back Pain Compared to Blank or no Treatment

Compared with blank or no treatment, this meta-analysis found that acupuncture achieves moderately better outcomes in the treatment of chronic low back pain. This corresponds to the experience of acupuncturists and the mechanism of pain relief. An acutely painful stimulus can help to relieve ongoing pain, which is known as “counter irritation” (Manchikanti et al., 2009). This phenomenon explains why practicing acupuncture without an analgesic earns more benefits than with analgesia. Furthermore, the tangible analgesia induced by acupuncture is caused by a central neurohumoral mechanism with delayed-onset characteristics (Cao, 2002). This indicates that acupuncture has direct and indirect effects on the area from which the nociceptive input originated.
Acupuncture Effects are Produced by Nonspecific Effects Arising from Manipulation

Sham acupuncture provides the same sting as acupuncture but without the sense of “De Qi”. According to the theory of traditional Chinese medicine, acupuncture outside Meridian lines cannot disturb or harmonize the Yin and Yang (Guan and Liu, 2009). In other words, theoretically sham acupuncture should be of no analgesic or therapeutic benefit. In our meta-analysis, acupuncture does not achieve better outcomes in terms of pain relief, disability rehabilitation or spinal flexion improvement compared with sham acupuncture. This indicates that the partial analgesic effect of acupuncture does not derive from the stimulation of the meridians, namely “De Qi”, but from the skin manipulation. Considering the effect sizes determined by the meta-analysis of these two subgroups in terms of pain intensity, disability and spinal flexion, we can conclude that a majority of the analgesic effect of acupuncture arises from other factors than “De Qi” itself. Consequently, it is not difficult to understand why better qualified and experienced acupuncturists are not more successful in treating patients with chronic low back pain (Witt et al., 2010).

Acupuncture Should Be Used in Combination with Other Treatments

Although the results of our meta-analysis show the superiority of acupuncture over other treatments, we draw a very cautious conclusion that the therapeutic effect of acupuncture may change along with different interventions of the other treatment subgroup. The main indication of acupuncture is pain, and acupuncture, as an alternative treatment, has proven itself to be useful for treating pain. Even so, acupuncture is not as effective as some other treatments, such as massage (Chou, 2010) and baclofen (Zaringhalam et al., 2010). We therefore recommend the use of acupuncture in combination with other treatments.

Chronic Low Back Pain for More than Three Weeks or More than 12 Weeks

We found that the effect of acupuncture in treating chronic low back pain was significantly limited by the duration of the disease, with a much smaller effect observable when the definition of chronic low back pain encompassed those with pain for more than 12 weeks. This indicates that the effect of acupuncture declines with prolonged disease duration, and explains the phenomenon that acupuncture often obtains a better outcome when treating acute pain than chronic pain (van Tulder et al., 1999).

Limitations and Strengths of Our Meta-Analysis

Our study had several limitations. The principal limitation was the relatively finite outcome variables of the involved RCTs, as we chose only four indexes to merge and ignored other meaningful indexes, such as adverse events and back-specific functional status, as outlined in Deyo et al. (1998). Another important limitation was the types of other treatments included in the analysis. In the stratified analysis, the other treatments subgroup was highly heterogeneous as much clinical heterogeneity exists in the different interventions for chronic low back pain. We factitiously merged these, which undoubtedly resulted in a
synthetic effect of acupuncture compared with other treatments. As a result of resource limitations, we did not search other Asia databases besides Chinese publications; this may have reduced our estimates of the effect sizes because acupuncture RCTs published in Asia have been shown to be positive in most cases (Vickers et al., 1998).

A major strength of this review was that we used a methodological guideline (Chou, 2010) for systematic reviews of trials for treatments of neck and back pain to report our meta-analysis. Ideally, we would have liked to report step-by-step following this guideline; however, the included studies did not provide enough valid data to complete each part of the analysis required in the guidelines. Furthermore, the intervention of acupuncture was checked and evaluated by STRICTA (MacPherson et al., 2010). Considering that the reporting of acupuncture trials is often disordered, we recommend the use of STRICTA as a checklist for authors in this field.

**Conclusions**

The current data shows that acupuncture is effective in providing long-term relief of chronic low back pain, but this effect is produced by non-specific effects that arise from skin manipulation. Further research is needed to evaluate which the effects of acupuncture compared with other treatments. Moreover, acupuncture should be used in combination with other treatments in the treatment of chronic low back pain. In view of the often refractory nature of chronic low back pain, comprehensive treatment plans, which may involve acupuncture, are urgently needed.

**References**


Guidelines for lower back pain treatment ignored. For patients with low back pain, the result is higher risk for adverse results and higher costs. *Duke Med. Health News* 16: 3, 2010.


