Andrea D. Furlan, Japanese Acupuncture and Moxibustion, 2010; Vol.6(1): 37-44

The 2 nd JSAM International Symposium on Evidence-based Acupuncture

A Systematic Review of Acupuncture for Chronic Low-back Pain

Andrea D. Furlan
Associate Scientist, Institute for Work & Health
Editorial Board, Cochrane Back Review Group
Assistant Professor, Department of Medicine, University of Toronto
Physician, Division of Physical Medicine and Rehabilitation, Toronto Rehabilitation Institute

Abstract

The objective of this review was to assess the effectiveness of needle acupuncture for chronic low-back pain. Acupuncture was compared to no treatment, sham acupuncture and to other therapies.

For this review we used the search strategy recommended by the Cochrane Back Review Group on MEDLINE, EMBASE, and CENTRAL with no language restriction, up to July 2008. We only included randomized controlled trials. The quality appraisal was performed with the 11-item recommended by the Cochrane Back Review Group (Van Tulder et al, 2003).

Three trials (1 high and 2 low quality) showed that acupuncture was better than no treatment for both measures of pain and function. But these measures were taken only in the short-term.

Six trials (3 high and 3 low quality) showed no difference between acupuncture and sham acupuncture on both measures of pain and function. But two high quality trials showed some benefit of acupuncture over sham acupuncture.

Five trials (4 high and 1 low quality) compared acupuncture to various treatments (massage, self-care, conventional therapy, TENS and spinal manipulation) and they showed variable results.

Seven trials (five high and 2 low quality) showed consistently the benefits of adding acupuncture to other therapies, compared to the other therapies alone, which included mostly exercises and physiotherapy.

In conclusion,

- Acupuncture is better than no treatment
- There is inconclusive evidence against sham acupuncture - more studies are needed to demonstrate benefits beyond placebo
- Acupuncture is no better than other treatments
- There is consistent evidence for the addition of acupuncture to other therapies

Key words: *Acupuncture, low back pain, randomized controlled trial, efficacy, effectiveness*

Introduction

Acupuncture is a millenary therapeutic technique developed in Asia and used for treatment and cure of many, if not all, diseases. Acupuncture has evolved and has endured as a reliable technique for its good results. Not until 50 years ago, acupuncture had not been known in many Western cultures. Not surprisingly, the popularity of acupuncture has grown among the Western civilizations and it has gained respect by many authorities in the health care field.

Corresponding author: Andrea D. Furlan
Institute for Work & Health, 481 University Avenue Suite 800, Tronto, ON M 5 G 2 E 9, Canada
E-mail: afurlan@iwh.on.ca

Concomitantly to the entry of acupuncture in the Western world was the introduction of a culture of questioning the effectiveness of all therapeutic modalities. The methodology of randomized controlled trials was initiated and rapidly improved. With the explosion of publications of randomized trials, the Cochrane Collaboration was founded in 1993 with the aim of preparing, maintaining and promoting systematic reviews of randomized trials in all areas of medicine. There are now 52 review groups all over the world responsible for the production of systematic reviews of health care interventions.

The Cochrane Collaboration Back Review Group (CBRG) is one of 52 Review Groups that produces and maintains systematic reviews on the effects of healthcare interventions. The scope of the CBRG is primary and secondary prevention and treatment of neck and back pain and other spinal disorders, excluding inflammatory diseases and fractures.

The CBRG published 40 reviews and 11 protocols (reviews in progress) in The Cochrane Library 2009, issue 3, released July 8th , 2009. It is the policy of The Cochrane Collaboration to update reviews every two years and to withdraw them if they are out of date. Review teams are currently updating a number of reviews that we expect to be published over the next few months.

The Cochrane Collaboration released an update of the *Handbook for Systematic Reviews of Interventions*[1)] and Review Manager 5, the software used for producing reviews (March 2008) and the CBRG Editorial Board finished their Updated Methods for Systematic Review in the Cochrane Back Review Group[2)]. There are several new features in the *Handbook*, the *Updated Guidelines* and the software that have been developed to make our reviews more transparent and user friendly.

Updated Methods Guidelines for Systematic Reviews in the Cochrane Back Review Group

It is recognized that differences in designs and conduct of individual studies can impact the validity (risk of over or underestimation of the true intervention effect) and rigour of the findings and final conclusion of the systematic review. Assessing the risk of bias (internal validity) of included studies is an integral step in systematic reviews. For over a decade, the CBRG has recommended the use of 11 criteria to assess the risk of bias in primary studies. The new Cochrane Handbook recommends one that has not previously been considered, bringing the currently recommended number to 12. The criteria are described below:

1. Was the method of randomization adequate?
2. Was the treatment allocation concealed?
3. Was the patient blinded to the intervention?
4. Was the care provider blinded to the intervention?
5. Was the outcome assessor blinded to the intervention?
6. Was the drop-out rate described and acceptable?
7. Were all randomized participants analysed in the group to which they were allocated?
8. Are reports of the study free of suggestion of selective outcome reporting? (NEW ITEM ADDED IN 2009)
9. Were the groups similar at baseline regarding the most important prognostic indicators?
10. Were co-interventions avoided or similar?
11. Was the compliance acceptable in all groups?
12. Was the timing of the outcome assessment similar in all groups?

The CBRG has found that if more than six criteria are met and there are no serious threats to the internal validity of the study, for example, over 50% of the participants are lost to follow-up, a study can be assessed as having a low risk of bias[3)].

Because of the high risk of bias and the heterogeneity of important components in many trials in this field, the CBRG has used Levels of Evidence to help synthesis the results, first used in the 1994 Guidelines for the Management of Acute Low-Back Pain in Adults. Following a new approach introduced in the updated Cochrane Handbook, the Updated Method Guidelines now recommend that the overall quality of the evidence for each outcome be assessed by using an adapted GRADE approach[2,4)].

Five domains are considered in the assessment of the GRADE of the evidence: limitations of the study design, inconsistency, indirectness (inability to generalize) and imprecision (insufficient or imprecise data) of results and publication bias across all studies that measure that particular outcome.

Inconsistency refers to the lack of similarity of estimates of treatment effects for the outcome across studies. Study results are considered consistent when direction, effect size and statistical significance are sufficiently similar to lead to the same conclusions. Consistency in direction is defined as 75% or more of the studies showing either a benefit or no benefit. In the case of a benefit, consistency in effect size is defined as 75% or more of the studies showing a clinically important or unimportant effect (see section on clinical relevance). Consistency in statistical significance is defined by the Chi test for heterogeneity.

Indirectness (lack of ability to generalize) refers to the extent to which the people, interventions and outcomes in the trials are not comparable to those defined in the inclusion criteria of the review. Authors may suggest that their results are more applicable to a specific population, (e.g. the effects of using insoles for young, male army recruits rather than a general working population) or that the results are based on an indirect comparison.

Imprecision refers to the number of participants and events and the width of the confidence interval for each outcome, especially when the confidence interval is sufficiently wide so that the estimate could either support or refute the effectiveness of the index intervention. Data are also imprecise when only one study reports an outcome, regardless of the sample size or the confidence interval and when fewer than 75% of the studies present data that can be included in a meta-analysis.

Publication bias refers to the probability of selective publication of trials and outcomes. This bias might be considered if full results for planned outcomes identified in a protocol or the trial report are not provided in the results section.

The quality starts at high when at least 75% of the RCTs with a low risk of bias provide consistent, direct, generalizable results for the outcome, and reduces by a level for each of the domains not met.

The overall quality of the evidence for each outcome is the result of the combination of the assessments in all domains. The GRADE Working Group recommends four levels of evidence:

- High quality evidence (☆☆☆☆) = at least 75% of the RCTs with no limitations of study design have consistent findings, direct and precise data and no known or suspected publication biases.
- Moderate quality evidence (☆☆☆) = one of the domains is not met
- Low quality evidence (☆☆) = two of the domains are not met
- Very low quality evidence (☆) = three of the domains are not met.
- No evidence = no RCTs were identified that addressed this outcome.

The CBRG welcomes consumers and experienced authors and referees to the group. We invite you to join on-line (www.cochrane.iwh.on.ca) to receive periodic newsletters that will keep you abreast of initiatives in The Cochrane Collaboration in general and the Cochrane Back Review Group in particular.

The Cochrane Review of Acupuncture for low-back pain

In 1999, the CBRG published a review of acupuncture for low-back pain.[5] It included 11 randomized controlled trials (RCTs), most of low methodological quality, and the authors refrained from making any firm conclusion because of the paucity of trials and their low quality.

In 2005, the CBRG updated this review and broadened the search strategies to Chinese and Japanese articles.[6,7] This review included 35 RCTs. 20 were published in English, 7 in Japanese, 5 in Chinese, and 1 each in Norwegian, Polish, and German. For chronic low back pain, there was evidence of pain relief and functional improvement for acupuncture compared to no treatment or sham therapy. These effects were only observed immediately after the end of the sessions and in short-term follow-up. There was also evidence that acupuncture, added to other conventional therapies relieves pain and improves function better than the conventional therapies alone. However, the effects were only small.

In 2009, the CBRG is preparing an update of this review. The search strategies continued to include Chinese and Japanese articles, but this time it was broadened to Korean literature too. The methods for critical appraisal

of the trials have been modified to reflect the current recommendations in the 2009 Updated Method Guidelines described above. The syntheses have incorporated the GRADE system to draw conclusion[4].

For effect sizes we used the following classification as recommended by the CBRG[2]:

Small
- WMD less than 10% of the scale (e.g. <10 mm on a 100 mm VAS).
- SMD or "d" scores <0.5.

Medium
- MD 10 to 20% of the scale.
- SMD or "d" scores from 0.5 to < 0.8.

Large
- MD >20% of the scale.
- SMD or "d" scores 0.8.

Results

The updated searches to May 2009 have found 30 additional RCTs. Of these, 3 dealing with acute low-back pain, one of trigger point acupuncture for myofascial pain syndrome, 13 of unknown duration of low-back pain, and 13 of chronic low-back pain. Of the 13 RCTs of chronic low-back pain, one compared acupuncture to no treatment[8], four to sham intervention[8-11], three compared acupuncture to another intervention[12-15], one compared the additional of acupuncture to usual care[16], and five were set to compare different techniques of acupuncture for chronic low-back pain[17-21].

Discussion

According to these results, acupuncture may be useful as either a unique therapy for chronic low back pain or

Table 1. Is acupuncture better than no treatment (waiting list) for chronic low- back pain?

Number of trials	Limitations (risk of bias)	Inconsistency (heterogeneity)	Directness (generalizability)	Imprecision (sparse data)	Summary of findings				GRADE
					N Acup	N control	Effect size		
Pain intensity - immediately after									
1	No serious limitations	No serious inconsistency[1]	No serious indirectness	Serious imprecision[2]	140	74	-0.88 (-1.17; -0.58)	large	High Moderate Low Very low No evidence
Pain intensity - short term (up to 3 months)									
2	Serious limitations[3]	No serious inconsistency[4]	No serious indirectness	Serious imprecision[5]	55	35	0.73 (-1.19; -0.28)	medium	High Moderate Low Very low No evidence
Pain intensity - intermediate term (3 - 12 months)									
1	Serious limitations[6]	No serious inconsistency	No serious indirectness	Serious imprecision[7]	30	10	-0.78 (-1.52; -0.04)	medium	High Moderate Low Very low No evidence
Pain intensity - long term (> 1 year)									
0								0	High Moderate Low Very low No evidence

[1] Only 1 trial
[2] Narrow CI and large sample size, but only 1 trial
[3] Both trials were judged to have high risk of bias
[4] No statistical heterogeneity (I^2=0%)
[5] Two trials, but sample size total 90.
[6] The trial was judged to have high risk of bias
[7] Large confidence interval, only 1 trial, and small sample size

There is "moderate grade" evidence of a large effect size in favour of acupuncture compared to no treatment immediately after the end of the sessions for pain reduction. These benefits are still maintained at the short and intermediate term follow-ups, but the effect sizes are medium and the grade of the evidence is "low".

Table 2. Is acupuncture better than sham/placebo for chronic low-back pain?

Number of trials	Limitations (risk of bias)	Inconsistency (heterogeneity)	Directness (generalizability)	Imprecision (sparse data)	Summary of findings N Acup	N control	Effect size		GRADE
Pain intensity - immediately after									
8	Serious[8]	No serious inconsistency[9]	Serious indirectness[10]	No serious imprecision[11]	705	645	0.30 (-0.44; -0.16)	small	High Moderate Low Very low No evidence
Pain intensity - short term (up to 3 months)									
3	No serious limitations	Serious heterogeneity[12]	No serious indirectness	Serious imprecision[13]	454	433	-0.49 (-0.1; 0.03)	No difference	High Moderate Low Very low No evidence
Pain intensity - intermediate term (3 - 12 months)									
4	No serious limitation	No serious heterogeneity	No serious indirectness	Serious imprecision[14]	573	486	-0.15 (-0.27; -0.03)	very small	High Moderate Low Very low No evidence
Pain intensity - long term (> 1 year)									
2	No serious limitation	No serious heterogeneity	No serious indirectness	Serious imprecision[15]	158	74	-0.21 (-0.49; 0.06)	No difference	High Moderate Low Very low No evidence

[8] 6 trials with low ROB and 2 trials with high ROB
[9] No statistical heterogeneity (I^2=18%)
[10] One trial (Inoue 2006 used a single session of acupuncture for CLBP)
[11] 8 trials, narrow confidence interval
[12] Statistical heterogeneity (I^2=82%)
[13] Wide confidence interval
[14] Narrow confidence interval, but very small effect size (-0.15)
[15] Wide confidence interval, small sample size in the control group (n=74)

For all follow-up times, there is either a small (but statistically significant) or a non-significant difference between real acupuncture and sham/placebo intervention. The grade of evidence varied from "low" to "moderate".

Table 3. Is acupuncture better than other conservative therapies for chronic low-back pain?

Studies	Risk of Bias	Comparison treatment	Immediate Up to 1 w	Short-term Between 1w-3m	Inter-mediate Between 3m -1 year	Long term ≥ 1 year
Cherkin 2001	low	Massage	-			<
Cherkin 2001	low	Self-care	-			-
Lehmann 1986	high	TENS	-		-	
Tsukayama 2002	low	TENS	+			
Muller 2003	low	SMT		<		
Li 2005	high	Electrotherapy	+		+	
Thomas 2006	low	Usual care				+
Zeng 2005	Iligh	Electrotherapy ultrashort-wave	+			

+ denotes acupuncture more effective; - denotes no difference in effectiveness; < denotes acupuncture less effective
TENS= transcutaneous electrical nerve stimulation, RTW = return to work

There is mixed evidence regarding acupuncture and other treatments Acupuncture is better than electrotherapy, and usual care alone. Acupuncture is not better than massage, spinal manipulation and self-care education. There is conflicting evidence regarding acupuncture compared to TENS.

Table 4. Is the addition of acupuncture to other intervention better than the other intervention alone?

Number of trials	Limitations (risk of bias)	Inconsistency (heterogeneity)	Directness (generalizability)	Imprecision (sparse data)	Summary of findings				GRADE
					N Acup	N control	Effect size		
Pain intensity - immediately after									
4	No serious limitation	No serious inconsistency	No serious indirectness	No serious imprecision	143	146	-0.76 (-1.02; -0.5)	medium	High Moderate Low Very low No evidence
Pain intensity – short term (up to 3 months)									
3	No serious limitation	Serious inconsistency[16]	No serious indirectness	No serious imprecision	97	85	-1.1 (-1.62; -0.58)	large	High Moderate Low Very low No evidence
Pain intensity – intermediate term (3 – 12 months)									
2	No serious limitation	No serious inconsistency	No serious indirectness	Serious imprecision[17]	59	56	-0.76 (-1.14; -0.38)	medium	High Moderate Low Very low No evidence
Pain intensity – long term (> 1 year)									
0									High Moderate Low Very low No evidence

[16] Statistical heterogeneity (I^2=62%)
[17] Only two trials with total sample size of 115

There is consistent "high" or "moderate grade" of evidence in favour of adding acupuncture to other interventions. The effect sizes were from medium to large.

Table 5 shows the summary of the GRADEs of the evidence and the Summary of Findings (SOF).

	Immediately after		Short-term		Intermediate-term		Long-term	
	GRADE	SOF	GRADE	SOF	GRADE	SOF	GRADE	SOF
Acupuncture compared to no treatment	☆☆☆	1 RCT Large ES	☆☆	2 RCTs Medium ES	☆☆	1 RCT Medium ES	O	O
Acupuncture compared to sham/placebo	☆☆	8 RCTs Small ES	☆☆	3 RCTs No diff	☆☆☆	4 RCTs Very small ES	☆☆☆	2 RCTs No diff
Acupuncture added to other intervention compared to the other intervention alone	☆☆☆☆	4 RCTs Medium ES	☆☆☆	3 RCTs Large ES	☆☆☆	2 RCTs Medium ES	O	O

GRADE of the evidence: O no evidence; ☆ Very low; ☆☆ Low; ☆☆☆ Moderate; ☆☆☆☆ High
SOF: Summary of Findings; ES: Effect Size; RCT: Randomized controlled trial

as an adjunct therapy to other conventional therapies. The most intriguing finding is the small difference between real acupuncture and sham/placebo acupuncture. Our review included 10 trials of sham interventions. One study (Lehmann 1986) was not included in the meta-analysis because the way the data was reported. Six trials em ployed superficial needling at non-acupuncture points (Mendelson 1983, Leibing 2002, Molsberger 2002, Brinkhaus 2006, Haake 2007, and Kown 2007). Two trials used sham-TENS (Carlsson 2001, and Kerr 2003), and one trial used a non-penetrating stimulation with the guide tube touching the skin at the most painful spot (Inoue 2006).

The use of superficial needling at non-acupuncture points has been criticized by not being a perfect sham intervention, due to the potential analgesic stimulation. More recently, there has been the development of sham

acupuncture devices that can blind not only the patient but the practitioner. One of these devices was developed by Takakura and Yajima[22,23]. This non-penetrating placebo needle has the needle tip simply pressing against the skin, and a matched penetrating needle. The needles are encased inside an opaque guide tube and the appearance and feel of the pair are designed to be indistinguishable. Sham needles, such as the Streitberger needle, employ a blunt tip which recedes into a hollow shaft when pressed against the skin thus, simulating penetration[24].

In 2008, Madsen et al published an analysis of 13 trials of acupuncture for various kinds of pains, with the objective to study the analgesic effect of acupuncture and placebo acupuncture and to explore whether the type of the placebo acupuncture is associated with the estimated effect of acupuncture. They found a small difference between real and sham acupuncture equivalent to 4 mm on a scale from 0 to 100 mm VAS. No association was detected between the type of placebo acupuncture and the effect of acupuncture[25].

Our findings are similar to the recent systematic review published by Yuan in 2008[26]. They included 23 randomized trials and found moderate evidence that acupuncture is more effective than no treatment, and strong evidence of no significant difference between real acupuncture and sham acupuncture for short-term pain relief. They found strong evidence that acupuncture can be a useful supplement to other forms of conventional therapy for nonspecific low-back pain, but the effectiveness of acupuncture compared with other forms of conventional therapies still requires further investigation.

References

1) Higgins,J.a.G.S.e. Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.0 [updated February 2008]. The Cochrane Collaboration, (2008).
2) Furlan,A., Pennick,V., Van Tulder, M., Bombardier, C. & The Editorial Board of the Cochrane Back Review Group. 2009 Updated Method Guideline for Systematic Reviews in the Cochrane Back Review Group. Spine. 2009.
3) Van Tulder, M.W., Suttorp, M., Morton, S., Bouter, L.M. & Shekelle, P.G. Empirical Evidence of an Association between Internal Validity and Effect Size in Randomized Controlled Trials. Spine. 2008.
4) Atkins, D., Best, D. & Briss, P.A. Grading quality of evidence and strength of recommendations. BMJ 328, 1490 (2004).
5) Van Tulder, M.W., Cherkin, D., Berman, B., Lao, L. & Koes, B.W. Acupuncture for low back pain. Cochrane Database Syst Rev. (1999).
6) Furlan, A.D. et al. Acupuncture and dry-needling for low back pain: an updated systematic review within the framework of the cochrane collaboration. Spine 30, 944-63 (2005).
7) Furlan, A.D. et al. Acupuncture and dry-needling for low back pain. Cochrane Database Syst. Rev. CD 001351 (2005).
8) Brinkhaus, B. et al. Acupuncture in patients with chronic low back pain: a randomized controlled trial. Arch Int Med 166, 450-7 (2006).
9) Haake, M. et al. German Acupuncture Trials (GERAC) for chronic low back pain: randomized, multicenter, blinded, parallel-group trial with 3 groups. Arch Int Med 167, 1892-8 (2007).
10) Inoue, M. et al. Relief of low back pain immediately after acupuncture treatment--a randomised, placebo controlled trial. Acupuncture in medicine : journal of the British Medical Acupuncture Society 24, 103-8 (2006).
11) Kown, Y.D. et al. The short-term efficacy of acupuncture for chronic low back pain: randomized sham controlled trial.[Korean]. J Orient Rehab Med 17, 123-32 (2007).
12) Li, N., Wu, B. & Wang, C.W. Comparison of acupuncture-moxibustion and physiotherapy in treating chronic non-specific low back pain. Chinese Journal of Clinical Rehabilitation 9, 186-7 (2005).
13) MacPherson, H., Thorpe, L., Thomas, K. & Campbell, M. Acupuncture for low back pain: traditional diagnosis and treatment of 148 patients in a clinical trial. Comp Ther Med 12, 38-44 (2004).
14) Thomas, K.J. et al. Randomised controlled trial of a short course of traditional acupuncture compared with usual care for persistent non-specific low back pain. BMJ (Clinical research ed.) 333, 623 (2006).
15) Zeng, H.-H., Wu, S.-L., Huang, L.-R., Ma, C. & Xiao, X.-H. Effect of acupuncture combined with cupping on the pain index in patients with chronic lumbar muscle strain. [Chinese]. Zhongguo Linchuang Kangfu 9, 46-7 (2005).

16) Witt, C.M. et al. Pragmatic randomized trial evaluating the clinical and economic effectiveness of acupuncture for chronic low back pain. American Journal of Epidemiology 164, 487-96 (2006).
17) Itoh, K., Katsumi, Y. & Kitakoji, H. Trigger point acupuncture treatment of chronic low back pain in elderly patients--a blinded RCT. Acupunct Med 22, 170-7 (2004).
18) Itoh, K., Katsumi, Y., Hirota, S. & Kitakoji, H. Effects of trigger point acupuncture on chronic low back pain in elderly patients - A sham-controlled randomised trial. Acupunct Med 24, 5-12 (2006).
19) Sator-Katzenschlager, S.M. et al. The short- and long-term benefit in chronic low back pain through adjuvant electrical versus manual auricular acupuncture.[see comment]. Anesthesia & Analgesia 98, 1359-64 (2004).
20) Tsui, M.L. & Cheing, G.L. The effectiveness of electroacupuncture versus electrical heat acupuncture in the management of chronic low-back pain. Journal of Alternative & Complementary Medicine 10, 803-9 (2004).
21) Yuan, J., Purepong, N., Kerr, D., Bradbury, I. & McDonough, S. A pilot randomised controlled trial comparing different frequencies of acupuncture for chronic low back pain. Focus on Alternative and Complementary Therapies 11, 57-8 (2006).
22) Takakura, N. & Yajima, H. A placebo acupuncture needle with potential for double blinding - a validation study. Acupunct. Med 26, 224-30 (2008).
23) Takakura, N. & Yajima, H. A double-blind placebo needle for acupuncture research. BMC Complement Altern. Med 7, 31 (2007).
24) Streitberger, K. & Kleinhenz, J. Introducing a placebo needle into acupuncture research. Lancet 352, 364-5 (1998).
25) Madsen, M.V., Gotzsche, P.C. & Hrobjartsson, A. Acupuncture treatment for pain: systematic review of randomised clinical trials with acupuncture, placebo acupuncture, and no acupuncture groups. BMJ 338, a 3115 (2009).
26) Yuan, J. et al. Effectiveness of acupuncture for low back pain: a systematic review. Spine 33, E 887-E 900 (2008).