Pain 96 (2002) 189–196

PAIN

www.elsevier.com/locate/pain

Acupuncture treatment of chronic low-back pain – a randomized, blinded, placebo-controlled trial with 9-month follow-up

Eric Leibing[a,*], Urs Leonhardt[b], Georg Köster[c], Anke Goerlitz[d], Joerg-André Rosenfeldt[c], Reinhard Hilgers[d], Giuliano Ramadori[b]

[a]*Department of Psychosomatics and Psychotherapy, Georg-August-University Goettingen, von-Siebold-Strasse 5, 37075 Goettingen, Germany*
[b]*Medical Clinic, University Clinic Goettingen, Robert-Koch-Str. 40, 37075 Goettingen, Germany*
[c]*Department of Orthopaedics, University Clinic Goettingen, Robert-Koch-Str. 40, 37075 Goettingen, Germany*
[d]*Department of Medical Statistics, Biometrical Study Center, Georg-August-University Goettingen, Humboldtallee 32, 37073 Goettingen, Germany*

Received 31 July 2001; received in revised form 26 September 2001; accepted 8 October 2001

Abstract

There is some evidence for the efficacy of acupuncture in chronic low-back pain (LBP), but it remains unclear whether acupuncture is superior to placebo. In a randomized, blinded, placebo-controlled trial, we evaluated the effect of traditional acupuncture in chronic LBP. A total of 131 consecutive out-patients of the Department of Orthopaedics, University Goettingen, Germany, (age = 48.1 years, 58.5% female, duration of pain: 9.6 years) with non-radiating LBP for at least 6 months and a normal neurological examination were randomized to one of three groups over 12 weeks. Each group received active physiotherapy over 12 weeks. The control group ($n = 46$) received no further treatment, the acupuncture group ($n = 40$) received 20 sessions of traditional acupuncture and the sham-acupuncture group ($n = 45$) 20 sessions of minimal acupuncture.

Changes from baseline to the end of treatment and to 9-month follow-up were assessed in pain intensity and in pain disability, and secondary in psychological distress and in spine flexion, compared by intervention groups.

Acupuncture was superior to the control condition (physiotherapy) regarding pain intensity ($P = 0.000$), pain disability ($P = 0.000$), and psychological distress ($P = 0.020$) at the end of treatment. Compared to sham-acupuncture, acupuncture reduced psychological distress ($P = 0.040$) only. At 9-month follow-up, the superiority of acupuncture compared to the control condition became less and acupuncture was not different to sham-acupuncture.

We found a significant improvement by traditional acupuncture in chronic LBP compared to routine care (physiotherapy) but not compared to sham-acupuncture. The trial demonstrated a placebo effect of traditional acupuncture in chronic LBP.

Keywords: Acupuncture; Chronic low-back pain; Placebo; Randomized-controlled trial

1. Introduction

Chronic low-back pain (LBP) is an important health issue in Western countries and is associated with high medical expenses, disability, and inability to work (Van Tulder et al., 1995). There are many therapeutic strategies, but long-term effects of single therapeutic approaches remain limited (Van Tulder et al., 1997). Multidisciplinary pain programs have proven to be an effective option with long-term benefit (Flor et al., 1992; Pfingsten et al., 1997, Van Tulder et al., 2000b), but gained only limited acceptance in routine care. LBP is a frequent reason for using unconventional therapies, especially acupuncture (Eisenberg et al., 1997; Paramore, 1997). However, the efficacy of acupuncture remains controversial because of insufficient empirical evidence and unclear underlying mechanisms (Ramsay et al., 1998).

In traditional Chinese medicine, disease is seen as a sign of imbalance between two complementary polarities (yin and yang). Needling the appropriate position of the 361 traditional acupuncture points (World Health Organisation, 1993) is proposed to rebalance the functioning by contacting the vital energy Qi. Western research has suggested other mechanisms (Lewith and Kenyon, 1984). For instance, the stimulation of A-delta afferents by appropriate needling (de qi) induces a cascade of endorphins and monoamines (Mayer, 2000a).

Current meta-analyses of randomized-controlled acupuncture trials provided some evidence for the efficacy of acupuncture in chronic LBP (Ter Riet et al., 1990; Ernst

* Corresponding author. Tel.: +49-551-396735; fax: +49-551-394592.
E-mail address: eleibin@gwdg.de (E. Leibing).

0304-3959/02/$20.00
PII: S0304-3959(01)00444-4

and White, 1998; Ramsay et al., 1998; Van Tulder et al., 1999; Ezzo et al., 2000; Mayer, 2000b; Smith et al., 2000; Van Tulder et al., 2000a). It remains unclear whether acupuncture is superior to placebo. The methodological quality of the few corresponding trials has been criticized, and high-quality trials have been requested. Insufficient empirical evidence might be due to problems in defining an appropriate placebo (Ramsay et al., 1998). Although there is no gold standard, minimal (sham) acupuncture seems to be the best control treatment. In minimal acupuncture, needles are inserted superficially at positions other than the conventional acupoints, thereby causing slight stimulation (Vincent and Lewith, 1995).

The present study therefore compares traditional acupuncture both to sham (minimal) acupuncture and to active physiotherapy.

2. Methods

2.1. Study design

The randomized, placebo-controlled, prospective study was conducted at the outpatient clinic of the Department of Orthopaedics, University Goettingen, Germany, from January 1996 to July 1998. It was performed according to the acupuncture guidelines of the World Health Organisation (World Health Organisation, 1995) and to the proposals of the National Institute of Health (Ramsay et al., 1998). It was approved by the local Ethics Committee.

One hundred and fifty out-patients met the entry criteria (non-radiating pain for more than 6 months) and were assigned to one of the three treatment groups (traditional acupuncture, sham-acupuncture, control) using a computer-based randomization process. Treatment records were kept separately from assessment records, so that the physician was unaware of the affiliation of the patients. Patients were blinded regarding traditional and sham-acupuncture. The clinician performing the acupuncture (Dr Li) had to know the group affiliation. The patients were advised to continue existing medication but not to commence any new analgesic or treatment.

2.2. Patient selection

A consecutive sample of $n = 150$ out-patients aged 18–65 years with non-radiating LBP for at least 6 months were included initially. Exclusion criteria were an abnormal neurological status, concomitant severe disease, psychiatric illness, current psychotherapy, pathological lumbosacral anterior-posterior and lateral X-rays (except for minor degenerative changes), rheumatic inflammatic disease, planned hospitalization, and refusal of participation. The 150 patients gave informed consent and were randomized. Nineteen patients had to be excluded after randomization, but prior to treatment. There were no significant differences between the groups ($P = 0.762$, Fisher's exact test, 3×3 table) regarding the reason for exclusion. Six patients withdrew their given consent (AG:3; SG:1; CG:2), in 11 patients (AG:5; SG:4; CG:2) exclusion criteria (e.g. painlessness, hospitalization, severe disease) appeared prior to treatment, and two patients relocated (AG:2). Thus, the final study sample included 131 patients.

2.3. Study procedures

At baseline (T0, week 0), all patients underwent a complete evaluation including a medical history, a routine physical examination, a comprehensive physical examination of the lumbar spine, and a neurological investigation. The physical assessments including evaluation of spine flexion and the other outcome variables were performed by a blinded physician (Dr Rosenfeldt) at baseline (T0, week 0), at the end of treatment (T1, week 12), and 9 months after the end of treatment (T2, 9-month follow-up, week 52). The data documentation was done by a collaborator (M. Koehler) of the Department of Medical Statistics, Biometrical Study Center.

2.4. Interventions

All patients received standardized active physiotherapy of 26 sessions (each 30 min) over 12 weeks. It was performed by trained physiotherapists according to the Bruegger-concept (Bruegger, 1990), in which nociceptive somato-motoric blocking effects were seen as causal factors in pain syndromes. The aim was to remove a muscle imbalance through special training of proper posture and motion.

The *control group* (CG) received active physiotherapy with no other treatment (routine care). The *acupuncture group* (AG) additionally received 20 sessions (each 30 min) of traditional and standardized acupuncture by an experienced Taiwanese physician (Dr Li) over 12 weeks. Dr Chien-Kang Li obtained his degrees at the University for Chinese Culture, Taipeh (Taiwan), and at the University of Goettingen, Germany. In the first 2 weeks of treatment, acupuncture was done five times a week, and in the next 10 weeks once a week.

Acupuncture was performed as combined traditional body- and ear-acupuncture (Stux and Pomeranz, 1990). No additional intervention or conversation was done, except for a short explanation about the procedure. Patients were first treated supine for body-acupuncture, and then seated for ear-acupuncture. Twenty fixed body acupoints (nine bilateral, two single points) and six on the ear (alternately on one ear) were selected according to their function in traditional Chinese medicine (Stux and Pomeranz, 1990; World Health Organisation, 1993) and were needled in every patient (Table 1). We did not conduct a diagnostic procedure to determine individual acupoints.

Depth of body needling depended on the location of the acupoints and upon the body build (min to max 10–30 mm). Needles were correctly inserted and manually stimulated at the body points until the 'de qi' sensation of heaviness and

Table 1
Selection of acupoints

Body-points[a]		Ear-points[a]
Yaoyangguan (GV 3)[b]	Weizhong (BL 40)[c]	Os sacrum (38)
Mingmen (GV 4)[b]	Kunlun (BL 60)[c]	Parasympathicus (51)
Shenshu (BL 23)[c]	Yanglingquan (GB 34)[c]	Nervus ischiadicus (52)
Dachangshu (BL 25)[c]	Sanyinjiao (SP 6)[c]	Lumbosacrum (54)
Shangliao (BL 31)[c]	Yautungdien (extra meridian, at the back of the hand)[c]	Shenmen (55)
Ciliao (BL 32)[c]		Kidney (95)

[a] Nomenclature according to WHO (8).
[b] Single points.
[c] Bilateral points.

numbness was elicited, lasting 5–20 s. Body needles were left in situ for 30 min, ear needles were not stimulated and left in situ for 1 week. Wrapped one-way, stainless steel, sterilized needles, produced by Shenzou, Suzhan Hua Tuo, China were used. Body needles were 0.3 mm in thickness and 40 mm in length, ear needles 0.23 mm in thickness and of standardized size (small, medium). Ear needles were ring-shaped and fixed by plaster.

The *sham-acupuncture group* (SG) received 20 sessions (each 30 min) of minimal acupuncture by the same physician (Dr Li) over 12 weeks. Sham-acupuncture was done following the standards of minimal acupuncture (Vincent and Lewith, 1995). Needles were inserted superficially, 10–20 mm distant to the verum-acupoints, outside the meridians, and were not stimulated (no 'de qi').

Each acupuncture procedure (location of the acupoints and duration of needling) was documented to look for the compliance to the acupuncture protocol.

2.5. Efficacy

Primary outcome measures were pain intensity, quantified by a 10 cm visual analog scale (VAS-P, Scott and Huskisson, 1976) and pain disabilty, measured by the pain disability index (PDI, Tait et al., 1990). The total score consisting of seven areas of activity (min to max 0–70) measures the overall level of disability using numeric rating scales (0 = no disability; 10 = total disability).

Secondary outcome measures were psychological distress and spine flexion. Psychological distress was measured by the hospital anxiety and depression scale (Zigmond and Snaith, 1983), a 14 item instrument for use in non-psychiatric medical patients. The total score (range 0–42) is a measure of psychological distress. Because there is a strong relationship between psychological distress and chronic pain, this is a relevant outcome parameter in the treatment of chronic pain patients (Craig, 1999). Spine flexion was measured by fingertip-to-floor distance (min 0 cm).

2.6. Safety

Safety was evaluated on the basis of the physical examinations and the patient-reported study events.

2.7. Statistical analysis

The power calculation was performed for the overall tests of efficacy (*F*-test, analysis of variance (ANOVA)). According to the tables of Cohen (1988; table 8.3.13, p. 314) a total sample size of $N = 144$ is needed to detect a medium effect-size (ES) ($f = 0.30$) with a power of 90% at $\alpha = 0.05$. This ES would be represented (Cohen, 1988; p. 276–278) by a difference between the three groups to detect $m = 1.5$ cm regarding the outcome variable VAS-P (average standard deviation (SD) = 2.0), and $m = 8.0$ regarding the outcome variable PDI (average SD = 11.0).

A last-observation-carried-forward (LOCF) strategy was used in the data analysis for the drop-outs. In the LOCF analysis, the last value of a patient who dropt out was carried forward into all subsequent missing timepoints. Potential biases by this method are discussed later (restrictions). Data analyses were performed using the statistical program package SAS (SAS, 1996). All tests were two-sided and the level of statistical significance was set at $\alpha = 0.05$. *P* values were reported. Categorial data were analyzed using Fisher's exact test. Differences between groups at baseline and the overall differences between the changes within the groups were analyzed using an ANOVA for each outcome measure. When a significant overall effect was observed, post hoc comparison between the groups (AG–SG, AG–CG) was performed using the Tukey studentized range test at a procedurewise rate of 0.05.

Empirical ES as a standardized unit of change between groups were calculated by subtracting the mean of change of the CG from the mean of change of the experimental (acupuncture) group, and then dividing the resulting difference by the SD of the CG change (Cohen, 1988; Flor et al., 1992).

3. Results

3.1. Patient characteristics

Results are presented for the randomized 131 patients who started treatment. No differences were found between the three groups regarding the evaluated variables at base-

Table 2
Patient characteristics

	All patients	AG	SG	CG	Statistic	*P*
Sample size	131	40	45	46	–	–
Age, mean (SD), y	48.1 (9.7)	47.9 (11.1)	49.0 (9.4)	47.5 (8.9)	0.30[a]	0.738[a]
Female sex, no. (%)	76 (58.0%)	22 (55.0%)	27 (60.0%)	27 (58.7%)	0.23[b]	0.903[b]
Pain duration, mean (SD), y	9.6 (8.2)	8.7 (7.7)	9.5 (8.3)	10.6 (8.7)	0.55[a]	0.581[a]
Body-mass-index, mean (SD)	26.3 (4.4)	26.1 (4.0)	25.9 (3.8)	26.9 (5.2)	0.64[a]	0.526[a]
Back surgery in history, no. (%)	11 (8.4%)	4 (10.0%)	2 (4.4%)	5 (10.9%)	1.41[b]	0.530[b]
Analgesic medication, no. (%)	66 (50.4%)	24 (60.0%)	20 (44.4%)	22 (47.8%)	2.24[b]	0.340[b]
Married, no. (%)	100 (76.3%)	32 (80.0%)	36 (80.0%)	32 (69.6%)	1.80[b]	0.437[b]
Employed, no. (%)	107 (81.7%)	33 (82.5%)	35 (77.8%)	39 (84.8%)	0.77[b]	0.663[b]

[a] *F*-test, ANOVA.
[b] Fishers-exact test.

line including sex, age, and duration of pain (Table 2). The distribution of years of education corresponded to that of the general German population.

At baseline, there were also no significant differences between the groups (Table 3) regarding the outcome measures pain intensity, pain disability, and spine flexion. There was a significant difference regarding psychological distress with a higher score in the CG.

The rate of drop-outs (Fig. 1) was not significantly different between the groups (AG:7, SG:14, CG:16, 4×3 table, Fishers's exact test, $P = 0.176$).

Side-effects as drop-out reasons occurred in AG only (painfulness of acupuncture $N = 2$, problems with circulation during acupuncture $N = 1$), whereas ineffective intervention (persisting pain) was the reason for drop-out in CG and SG only (CG:5, SG:5). Painlessness was the reason for drop-out in two patients (AG:1, CG:1) and in three patients another severe disease or an accident occurred (SG:3). Two patients (SG:1, CG:1) wanted to change the group and one patient wanted to stop physiotherapy (AG:1). Ten patients failed to return (AG:01, SG:2, CG:7) and six patients (AG:1, SG:3, CG:2) requested to stop participation because of external reasons (job, relocation, expense).

3.2. Between group changes-efficacy

The course of the primary outcome criteria pain intensity (VAS-P) and pain disability (PDI) within the groups is visualized in Figs. 2 and 3.

The ANOVAs (Table 3) demonstrated significant overall differences between the within-group changes regarding pain intensity (T0−T1: $F = 7.47$, $P = 0.0009$), pain disability (T0−T1: $F = 10.97$, $P = 0.0001$; T0−T2: 4.98, $P = 0.0083$), and psychological distress (T0−T1: $F = 4.27$, $P = 0.0161$; T0−T2: $F = 2.03$, $P = 0.1354$), but not regarding spine flexion.

Table 3
Groups means (SD) of outcome measures at baseline and within-group changes

	AG $N = 40$	SG $N = 45$	CG $N = 46$	*F* value (*P* value)[a]
Primary outcome criterion: pain intensity (VAS-P)				
T0 Baseline	4.8 (1.8)	5.3 (1.8)	5.4 (1.9)	1.08 (0.342)
Δ T0−T1[b](pre−post)	−2.7 (2.2)	−2.1 (2.2)	−1.0 (1.7)	7.47 (0.0009)
Δ T0−T2[b](pre−follow-up)	−1.7 (1.8)	−1.8 (2.2)	−0.9 (2.0)	2.75 (0.068)
Primary outcome criterion: pain disability (PDI)				
T0 Baseline	25.2 (13.4)	25.5 (10.4)	24.9 (13.7)	0.03 (0.973)
Δ T0−T1[b](pre−post)	−13.9 (15.0)	−9.7 (10.5)	−2.6 (7.8)	10.97 (0.0001)
Δ T0−T2[b](pre−follow-up)	−9.0 (12.5)	−8.5 (11.3)	−2.3 (10.0)	4.98 (0.0083)
Secondary outcome criterion: psychological distress (HADS)				
T0 Baseline	12.7 (7.9)	12.5 (6.7)	16.2 (8.2)	3.45 (0.035)
Δ T0−T1[b](pre−post)	−4.0 (5.8)	−1.4 (4.7)	−1.2 (3.8)	4.27 (0.0161)
Δ T0−T2[b](pre−follow-up)	−3.6 (4.8)	−2.1 (5.2)	−1.3 (5.5)	2.03 (0.1354)
Secondary outcome criterion: spine flexion				
T0 Baseline	18.8 (14.5)	19.0 (12.1)	14.2 (13.7)	1.84 (0.163)
Δ T0−T1[b](pre−post)	−6.3 (15.0)	−2.8 (11.3)	−1.8 (10.7)	1.52 (0.223)
Δ T0−T2[b](pre−follow-up)	−4.9 (13.7)	−4.0 (11.2)	−0.2 (11.6)	1.87 (0.158)

[a] ANOVA, DF = 2, for details see statistics.
[b] Δ: differences between timepoints.

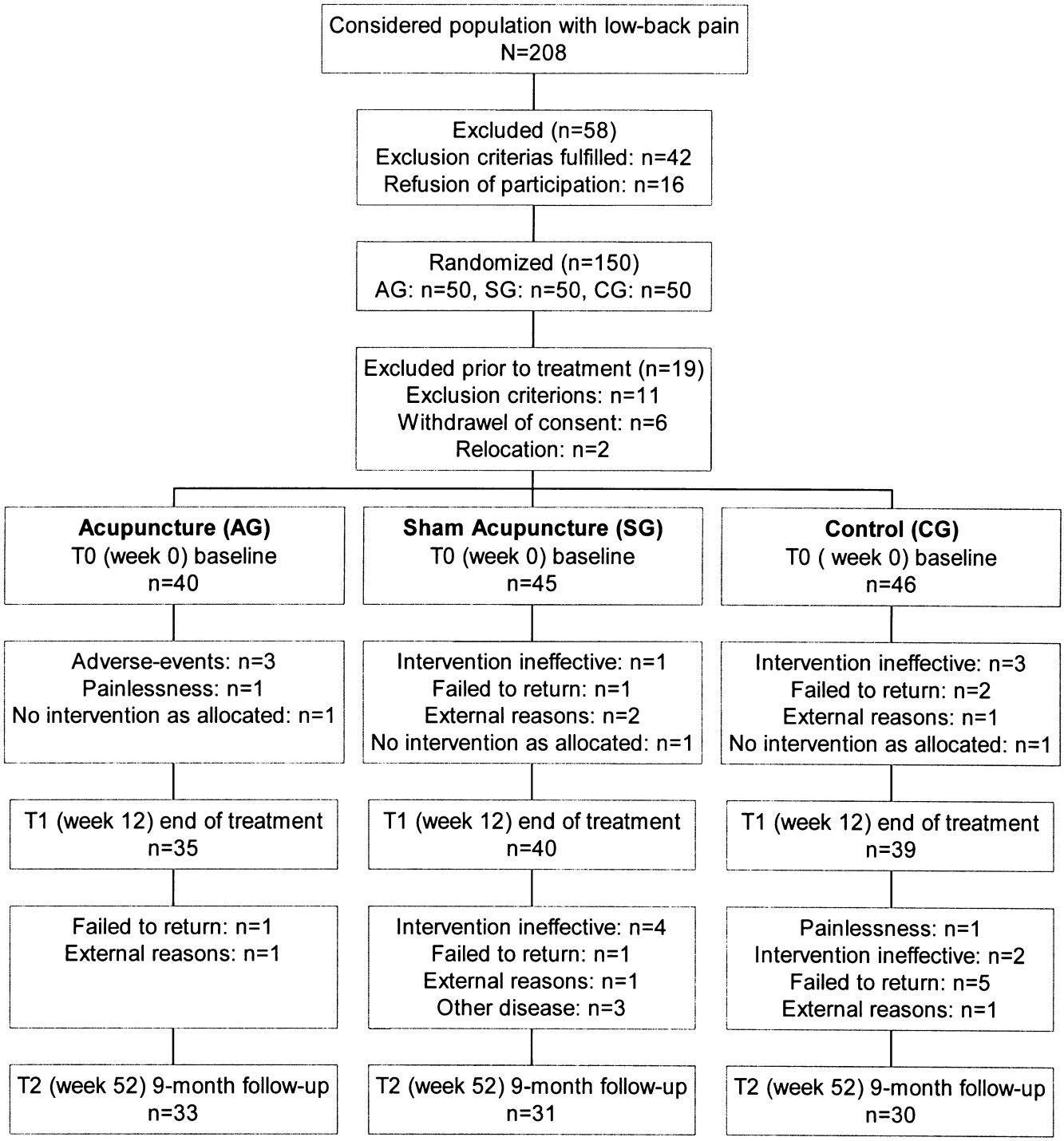

Fig. 1. Participation flow in the study.

Tukey's studentized range test was conducted to test for differences between the within-group changes of two groups (AG–SG, AG–CG) controlling for multiple testing. The differences of means (T0−T1; T0−T2) of the AG was compared to the differences of the other groups (Table 4).

At the end of treatment (T1, week 12) acupuncture was superior to the CG regarding pain intensity ($P = 0.000$, $ES = 1.00$), pain disability ($P = 0.000$, $ES = 1.45$), and psychological distress ($P = 0.020$, $ES = 0.73$). There was no significant difference in acupuncture and SG regarding pain intensity and pain disability. There was a significant difference between AG and SG with a significant reduction of psychological distress in the AG ($P = 0.040$, $ES = 0.55$).

Nine months after the end of treatment (T2, week 52) AG was still superior to CG regarding pain disability ($P = 0.016$,

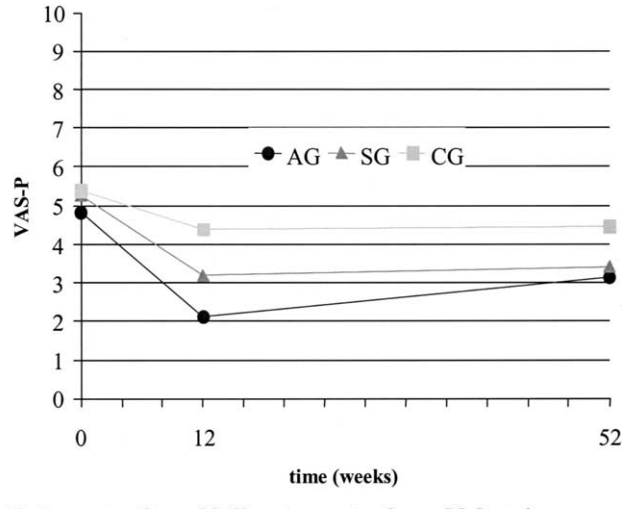

Fig. 2. Intensity of pain (VAS-P) over time.

$ES = 0.68$) only. In this period, there were no differences between AG and SG.

4. Discussion

In this randomized, placebo-controlled trial a sample of chronic LBP-patients characterized by high pain intensity, chronicity, and psychological distress was enrolled. Spontaneous improvement is unlikely because of these characteristics and the limited response to former treatments. Because there were no significant differences regarding sociodemographic and outcome variables except higher psychological distress in CG at baseline, successful randomization could be assumed. The exclusion of the 19 cases prior to treatment seems to be a reasonable strategy, because the efficacy of realized acupuncture is of main interest. Furthermore, the exclusion of these cases from analyses will not have led to a bias, because there are no significant differences regarding the reasons for exclusion between the groups.

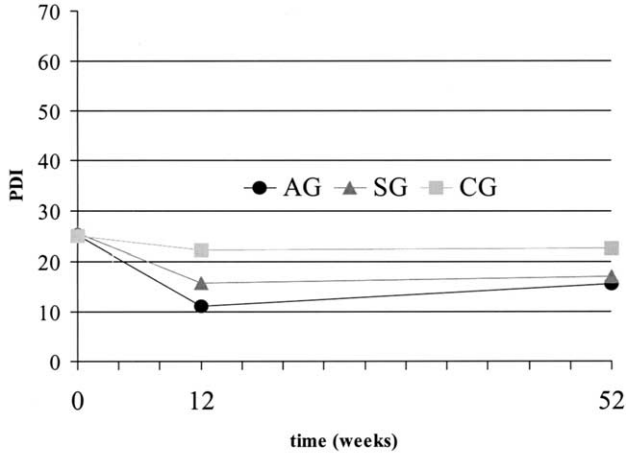

Fig. 3. Pain disability (PDI) over time.

The acupuncture was safe. Minor unserious adverse events occurred in three patients. However, possible serious adverse events in routine care may influence the risk/benefit relation (Ernst and White, 1997; McCartney et al., 2000).

4.1. Efficacy

We did not find any significant improvement in spine flexion, measured by fingertip-to-floor distance, but some authors argued a low reliability of this measure (Merritt et al., 1986).

In comparison to the CG (active physiotherapy), the primary outcome criteria pain intensity and pain disability were significantly improved, but we could find one significant difference (psychological distress) comparing traditional acupuncture to sham-acupuncture only. Furthermore, the differences between the groups were getting smaller over time.

Thus, the effect of acupuncture is not due to active physiotherapy alone. The effect of acupuncture might be unspecific or a placebo effect, if sham-acupuncture represents an appropriate control condition. These results are in accordance with several meta-analyses (Ter Riet et al., 1990; Ernst and White, 1998; Ramsay et al., 1998; Van Tulder et al., 1999; Ezzo et al., 2000; Mayer, 2000b; Smith et al., 2000; Van Tulder et al., 2000a), showing no or limited effects of verum acupuncture compared to placebo conditions in high-quality trials. An association of methodological quality and the estimation of a treatment effect has been postulated with the most valid trials that seem to indicate no or small effects of acupuncture compared to placebo (Ezzo et al., 2000; Mayer, 2000b; Smith et al., 2000).

4.2. Restrictions and potential biases

All patients agreed to study protocol, so the participation's expectations may have affected treatment response. By building up an SG, we tried to control for the possible influence of this variable. These groups were balanced also regarding time and attention. However, sham-acupuncture did not fulfill all criteria of an appropriate control condition, since specific elements of needling cannot be separated completely from non-specific treatment elements.

Minimal acupuncture (Vincent and Lewith, 1995) mimicked the traditional acupuncture technique and minimized afferent stimulation by superficial insertion of needles without stimulation and by needling distant to verum acupuncture points. It might be speculated that most participants expected pain by needling. Therefore, the elimination of any afferent stimulation, e.g. by using a local anaesthetic, may lead to an incomplete blinding of the participants.

The acupuncturist had to know the group assignment of the patients. To exclude a potential bias, the interaction with participants was minimized, and evaluation was done by an

Table 4
Changes in outcome measures between groups

Variable	Contrast[a]	ΔT0−T1[b](pre−post)			ΔT0−T2[b](pre−follow-up)		
		Difference (95% CI)[c]	P[c]	ES[c]	Difference (95% CI)[c]	P[c]	ES[c]
Pain intensity	AG–SG	−0.6 (−1.65 to 0.45)	0.366	0.27	−0.1 (−1.21 to 0.87)	0.921	−0.07
	AG–CG	−1.7 (−2.71 to −0.62)	0.000	1.00	−0.8 (−1.93 to 0.07)	0.186	0.38
Pain disability	AG–SG	−4.2 (−9.99 to 1.71)	0.214	0.39	−0.5 (−6.33 to 5.35)	0.978	−0.04
	AG–CG	−11.3 (−17.01 to −5.44)	0.000	1.45	−6.8 (−12.57 to −0.96)	0.016	0.68
Psychological distress	AG–SG	−2.6 (−5.02 to −0.06)	0.040	0.55	−1.5 (−4.15 to 1.20)	0.391	0.29
	AG–CG	−2.8 (−5.26 to −0.32)	0.020	0.73	−2.3 (−4.90 to 0.42)	0.114	0.42
Spine flexion	AG–SG	−3.5 (−9.88 to 2.88)	0.395	0.31	−0.9 (−7.09 to 5.43)	0.947	0.07
	AG–CG	−4.5 (−10.88 to 1.88)	0.216	0.42	−4.7 (−10.98 to 1.54)	0.174	0.41

[a] AG, acupuncture group; SG, sham-acupuncture group; CG, control group.
[b] Δ: differences between groups regarding changes between baseline (T0) and end of treatment (T1), and between baseline (T0) and 9-month follow-up (T2).
[c] Differences between the within-group changes, CI, 95% confidence-interval; P, empirical P-value; ES, effect-sizes (see statistics); negative scores indicating smaller reduction in AG.

independent physician and by questionnaires. The patients' stated expectations of pain relief did not correlate with real efficacy in acupuncture or placebo in former trials (Mendelson et al., 1983).

These possible psychological effects may have led to an over-estimation of the effects of verum acupuncture, but the possible small effects of afferent stimulation by superficial needling would lead to an underestimation of verum acupuncture, so these reversal effects may compensate each other.

In traditional Chinese medicine, the individualized selection of acupoints is an important principle. In contrast, the application of acupuncture at fixed acupoints is necessary in controlled trials. Therefore, the conclusions of our study are valid for the acupuncture at fixed acupoints only, and the outcome might be better with individualized treatment.

At the end of treatment (T1, week 12) the drop-out rate was 13% (17/131), but it increased in the follow-up period (T2, week 52) to 28.2% (37/131). There are no widely accepted references to 'acceptable' drop-out rates. The use of the LOCF is generally recommended as a method with no a priori implication of bias (Gilling and Koch, 1991). It may lead to biased estimates of outcome, particularly if drop-out rates differ among groups, and if pronounced changes in outcome variables are likely (Meyer and Windeler, 1998). With respect to the high level of pain intensity (VAS-P: 5.2 cm) and the chronicity (duration: 9.6 years) of our sample, spontaneous improvement or rapid worsening of pain seems unlikely. Furthermore, we could not find a significant difference between the drop-out rates in groups. Therefore, the use of the LOCF method seems to be an appropriate strategy to handle the drop-out problem in this trial.

5. Conclusions and proposals

There is no effect of traditional acupuncture beyond a placebo and probably a small unspecific effect of needling. But placebo analgesia has been proven to be effective and is also mediated by endogenous opiates (Ter Riet et al., 1990). Despite the uncertainty in the underlying mechanisms, and using the placebo and the unspecific effects, acupuncture may be nevertheless a viable option as an adjunct treatment in the pragmatic treatment of chronic LBP-patients (Ramsay et al., 1998). With an initial pain reduction, even caused by placebo mechanisms, general activation, and multimodal treatment programs (Van Tulder et al., 2000b) may be more likely to be accepted by chronic LBP-patients and may result in an improved outcome.

In future studies, efforts should be made to isolate the impact of individual components of acupuncture treatment like needling of different acupoints, kind of stimulation, and doses of acupuncture (Ezzo et al., 2000). This study did not test for psychological variables, which possibly influence the effect of acupuncture like hypnotic susceptibility or expectation of pain relief, but this may be interesting to look for in further studies, too.

Acknowledgements

This study was funded by the Ministry of Education, Science, Research and Technology (BMBFT), Federal Republic of Germany (01 KT 9407).

First, we thank C.-K. Li for conducting the acupuncture and for the critical comments. We thank M. Koehler for data documentation, V.W. Armstrong, F. Leichsenring, and S. Rabung for a critical review of the manuscript, and the patients for participation in our trial.

References

Bruegger A. Gesunde Koerperhaltung im Alltag, Zuerich: Bruegger Verlag, 1990.

Cohen J. Statistical Power Analysis for the Behavioral Sciences, 2nd ed. Hillsdale, NJ: Erlbaum, 1988.

Craig KD. Emotions and psychobiology. In: Wall PD, Melzack R, editors. Textbook of Pain, 4th ed. Edinburgh: Churchill Livingstone, 1999.

Eisenberg DM, Davis RB, Ettner SL, Appel S, Wilkey S, Van Rompay M, Kessler RC. Trends in alternative medicine use in the United States, 1990–1997; results of a follow-up national survey. J Am Med Assoc 1997;278:2111–2112.

Ernst E, White AR. Life-threatening adverse reactions after acupuncture? A systematic review. Pain 1997;71:123–126.

Ernst E, White AR. Acupuncture for back pain. A meta-analysis of randomized controlled trials. Arch Intern Med 1998;158:2235–2241.

Ezzo J, Berman B, Hadhazy VA, Jadad AR, Lao L, Singh BB. Is acupuncture effective for the treatment of chronic pain? A systematic review. Pain 2000;86:217–225.

Flor H, Fydrich T, Turk DC. Efficacy of multidisciplinary pain treatment centers: a meta-analytic review. Pain 1992;49:221–230.

Gilling D, Koch G. The application of the principle of intention-to-treat to the analysis of clinical trials. Drug Inf J 1991;25:411–424.

Lewith GT, Kenyon JN. Physiological and psychological explanations for the mechanism of acupuncture as a treatment for chronic pain. Soc Sci Med 1984;19:1367–1378.

Mayer DJ. Biological mechanisms of acupuncture. Prog Brain Res 2000a;122:457–477.

Mayer DJ. Acupuncture: an evidence-based review of clinical literature. Annu Rev Med 2000b;51:49–63.

McCartney CJL, Herriot R, Chambers WA. Bilateral hand oedema related to acupuncture. Pain 2000;84:429–430.

Mendelson G, Selwood TS, Kranz H, Loh TS, Kidson MA, Scott DS. Acupuncture treatment of chronic back pain. A double-blind placebo-controlled trial. Am J Med 1983;74:49–55.

Merritt JL, McLean TJ, Erickson RP, Offord KP. Measurement of trunk flexibility in normal subjects: reproducibility of three clinical methods. Mayo Clin Proc 1986;61:192–197.

Meyer K, Windeler J. Intention-to-treat in progressively deteriorating disease: should the last observation be carried forward? J Epidemiol Biostat 1998;3:199–208.

Paramore LC. Use of alternative therapies: estimates from the Robert Wood Johnson Foundation national access to care survey. J Pain Symptom Manage 1997;13:83–89.

Pfingsten M, Hildebrandt J, Leibing E, Franz C, Saur P. Effectiveness of a multimodal treatment program for chronic low-back pain. Pain 1997;73:77–85.

Ramsay DJ, Bowman MA, Greenman PE, Jiang SP, Kushi LH, Leeman S, Lin KM, Moerman DE, Schnoll SH, Walker M, Waternaux C, Wisneski LA. Acupuncture. NIH consensus development panel on acupuncture. J Am Med Assoc 1998;280:1518–1524.

SAS-Institute. SAS Language Reference. Version 6.12. Cary, NC: SAS-Institute, 1996.

Scott J, Huskisson EC. Graphic representation of pain. Pain 1976;2:175–184.

Smith LA, Oldman AD, McQuay HJ, Moore RA. Teasing apart quality and validity in systematic reviews: an example from acupuncture trials in chronic neck and back pain. Pain 2000;86:119–132.

Stux G, Pomeranz B. Basics of Acupuncture, 2nd ed. Berlin: Springer, 1990.

Tait RC, Chibnall JT, Krause S. The pain disability index: psychometric properties. Pain 1990;40:171–182.

Ter Riet G, Kleijnen J, Knipschild P. Acupuncture and chronic pain: a criteria-based meta-analysis. J Clin Epidemiol 1990;43:1191–1199.

Van Tulder MW, Koes BW, Bouter LM. A cost-of-illness study of back-pain in The Netherlands. Pain 1995;62:233–240.

Van Tulder MW, Koes BW, Bouter LM. Conservative treatment of acute and chronic nonspecific low-back pain: a systematic review of randomized controlled trials of the most common interventions. Spine 1997;18:2128–2156.

Van Tulder MW, Cherkin DC, Berman B, Lao L, Koes BW. The effectiveness of acupuncture in the management of acute and chronic low-back pain. Spine 1999;24:1113–1123.

Van Tulder MW, Cherkin DC, Berman B, Lao L, Koes BW. Acupuncture for low back pain. Cochrane Database Syst Rev 2000a;2:CD001351.

Van Tulder MW, Koes BW, Assendelft WJ, Bouter LM, Maljers LD, Driessen AP. Chronic low-back pain: exercise therapy, multidisciplinary programs, NSAID's, back schools and behavioral therapy effective; traction not effective; results of systematic reviews. Ned Tijdschr Geneeskd 2000b;44:1489–1494.

Vincent CA, Lewith G. Placebo-controls for acupuncture studies. J R Soc Med 1995;88:199–202.

World Health Organisation (WHO). Standard acupuncture nomenclature. A brief explanation of 361 classical acupuncture point names and their multilingual comparative list, 2nd ed. Manila (Regional Office for the Western Pacific): World Health Organisation, 1993.

World Health Organisation (WHO). Guidelines for clinical research on acupuncture, Manila (Regional Office for the Western Pacific): World Health Organisation, 1995.

Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand 1983;67:361–370.