A Bandolier paper

What do we think?
What do we know?
What can we prove?

Assessing the evidence of effectiveness of acupuncture for stroke rehabilitation: stepped assessment of likelihood of bias

Lesley A Smith, Owen A Moore, Henry J McQuay, Andrew Moore

Pain Research Unit and Nuffield Department of Anaesthetics
University of Oxford, Oxford Radcliffe Hospital
The Churchill, Headington, Oxford OX3 7LJ, UK

Correspondence to RA Moore, Pain Research Unit
Tel: +44 1865 226132 Fax: +44 1865 226978
email: andrew.moore@pru.ox.ac.uk

Abstract

Objective: To investigate the effects of known potential sources of bias on whether acupuncture is beneficial in treating stroke.

Data sources: Cochrane Library, MEDLINE, EMBASE, PubMed and reference lists of previous reviews were used to seek randomised controlled trials. There were seven randomised trials that met the inclusion criteria.

Methods: Pooling of outcomes was known to be unlikely. Trials were judged by the authors and by us as being positive (acupuncture was helpful) or negative (no benefit could be shown). The effect of trial quality, validity and origin were examined to investigate whether these affected the overall outcome. The overall outcome - is there evidence of a benefit and is the size of that benefit worthwhile? - was examined from the perspective of the highest quality trials.

Results: Our conclusion was that two studies had a positive result and five a negative. The three observer blind trials we judged to be negative. The single trial with a quality score of three was negative. The two trials with a validity score of nine or more we judged to be negative. Two of the three European studies we judged to be negative. Sensitivity analyses based on blinding, reporting quality, validity score and country of origin all showed a higher proportion of positive results for poor quality trials than for those of higher quality.

Conclusion: Sensitivity analysis for known sources of bias is important where "vote-counting" replaces data pooling. There were no high quality trials of acupuncture for stroke that showed that it was beneficial.

Introduction

Systematic reviews of alternative therapies are bedevilled by trials of poor quality. One result of this is that the reviews often conclude that the evidence for a therapy is inconclusive. The main issue is one of bias arising from inadequate trial design, choice of outcomes, data analysis or reporting.

There is good evidence that bias is a major problem. Schultz et al [1] demonstrated that lack of randomisation is the major source of bias in trials; studies which are not randomised can lead to overestimation of treatment effects by up to 40%. Restricting systematic reviews to include only randomised studies therefore makes sense for reviews of effectiveness. A classic example is a review of transcutaneous nerve stimulation (TENS) for postoperative pain relief. Randomised studies overwhelmingly showed no benefit over placebo, while non-randomised studies did show benefit [2].

Non-blinded studies over-estimate treatment effects by about 17% [1]. In a review of acupuncture for back pain [3], including both blinded and non-blinded studies changed the overall conclusion. The blinded studies showed 57% of patients improved with acupuncture and 50% with control, a relative benefit of 1.2 (95% confidence interval 0.9 to 1.5). Five non-blinded studies showed a difference from control, with 67% improved with acupuncture and 38% with control. Here the relative benefit was significant at 1.8 (1.3 to 2.4).

Trials of poor reporting quality consistently over estimate the effect of treatment. Using a scoring system for methodological quality [4], studies of lower quality are likely to over-estimate treatment effects [5,6]. Other sources of bias may include small trials [7-9], covert duplication [10], and geography [11]. Vickers and colleagues [11] showed that trials of acupuncture conducted in east Asia were universally positive, while those conducted in Australasia, north America or western Europe were positive only about half the time. Randomised trials of other therapies conducted in China, Taiwan, Japan or Russia/USSR were also overwhelmingly positive.

Then there is the issue of the overall validity of a randomised trial. In some areas, like acute pain, valid methods for the conduct of clinical studies have been set out for many years, and are well understood [12]. The result is many trials that are randomised and double blind, and conducted on patients with the same initial severity of pain under similar conditions and assessing identical or similar outcomes over the same time periods. Trials with low validity are more likely to have a positive result than those with higher validity [13], seen in acupuncture for head and neck pain. This mimics a finding for systematic reviews, where poor reporting quality of reviews also leads to a greater likelihood of a positive result [14,15].

Finally we need to know whether a technology works, and how well it works. "Vote-counting", where the number of positive and negative papers are added and compared, is particularly at risk from the influence of small studies with biased design. These are particularly prevalent for complementary and alternative therapies. The application of quality standards may change the overall impression of a technology, and this review set out to test that for the use of acupuncture following stroke.

Methods

Full published reports of randomised controlled trials (RCTs) of traditional and non-traditional acupuncture treatment for stroke rehabilitation were sought. Different search strategies were used to identify eligible reports in MEDLINE (1966 to January 2000), EMBASE (1980 to January 2000), CINAHL (1982 to 2000), PSYCHLIT (1982 to 2000) PubMed (July 2000), and the Cochrane Library (online July 2000). A broad free text search with no restriction to language was undertaken. Reference lists of retrieved reports and reviews [16] were searched for additional trials. Unpublished reports and abstracts were not considered. Authors were not contacted for original data.

Inclusion criteria were RCTs comparing acupuncture, with or without electrical stimulation, or laser acupuncture with a control group; patients had an acute stroke diagnosed by standard neurological tests; group size ≥10 and physical function outcomes.

Each report that could possibly be described as an RCT was read independently by three of the authors (LAS, OAM, RAM). Trials meeting inclusion criteria were screened independently and scored using a three item, 1-5 score, quality scale [4] and a 5 item, 0-16 score, validity scale [13]. From each trial data were extracted on trial design, acupuncture and control interventions, outcome measures, statistical analysis, and geographic location of the trial.

Results

There were seven studies with 505 patients that met the inclusion criteria [17-23]. Details of the studies are given in Table 1. Information from one trial [20] was reported three times [21, 24, 25]. We used the report with the longest outcomes. Two randomised studies were excluded: one measured sensory stimulation rather than traditional acupuncture [26], one measured the effect of acupuncture on microcirculation in the fingers [27].

All seven included studies examined the effects of acupuncture on patients following a first stroke, using a parallel group design, where acupuncture plus a standard treatment was compared with standard treatment alone. None used sham acupuncture as a control. In all cases electrical stimulation of the acupuncture needles was used, and three studies reported that the stimulation intensity was sufficient to cause muscle contraction [19,21,23]. The stimulation frequencies used in the studies ranged from 2 Hz to at least 25 Hz. Three studies were conducted in Scandinavia, and four in China or Taiwan.

A wide range of outcomes was described, and no study defined a primary outcome measure for effectiveness. Typically outcomes included some measure of motor function, some assessment of activities of daily living (Barthel's in-

dex), and frequently a patient assessed quality of life assessment (Nottingham Health Profile). No study was double blind. Three were single blind and used an observer blinded to the treatment given [20-22]. The other studies made no attempt to blind the observations.

Reporting quality overall was poor. Only one study reported the method used for randomisation [21]. Most reported patient withdrawals (Table 1). One study had a quality score of three [21] and all others were two or below. Validity scores were nine or above in two studies [21,22] and the others ranged from four to eight (Table 1).

In only one study [21] did the original authors conclude that acupuncture was ineffective. In four studies we did not agree with the authors' conclusions [18,20,22,23]. The reason for the difference (Table 1) was because the data presented in the paper did not support their conclusion. Typically some derivative index (change from baseline) was compared between acupuncture and control and found to be better with acupuncture, while there remained no difference in the absolute values at either baseline or at the time of assessment. One study [23] had results which were unbelievable, including identical control scores in 59 patients for seven different outcomes.

The full reasons for disagreement were:

- Hu et al [18]: positive result depended only on sub group analysis and on only some outcomes.
- Kjendahl et al [20]: statistical significance only obtained on small differences between start and finish points.
- Si et al [22]: there was no difference between patients at beginning and end of study.
- Wong et al [23]: statistical analysis and reported results are not believable. For instance Table 4 in the paper has exactly the same mean symptom improvement score and variability for seven different items.

Overall, our conclusion was that two studies had a positive result and five a negative (Table 2). The three observer blind trials we judged to be negative. The single trial with a quality score of three was negative. The two trials with a validity score of nine or more we judged to be negative. Two of the three European studies we judged to be negative.

Discussion

The problem faced by readers of medical publications and reviews is the credibility of the result. For clinical trials of treatment efficacy we now know that there are many sources of bias, all of which tend to overestimate the benefit of treatment. In this review we sought all reported trials of acupuncture for stroke rehabilitation if they met minimum requirements of randomisation, size, and clinical outcomes.

We found seven trials. There were major problems, not least the fact that their authors drew incredible conclusions from the data presented. We disagreed with the conclusions of four of seven studies. For acupuncture in neck and back pain disagreement was limited to two of 16 studies [13].

Performing a sensitivity analysis in any systematic review is not only sensible but should be mandatory [28]. In this review sensitivity analyses based on blinding, reporting quality, validity score and country of origin all showed a

Table 2: Outcome according to original authors and reviewers, and according to various sources of bias.

	Conclusion of original authors		Conclusion of reviewers	
Potential source of bias	**Positive**	**Negative**	**Positive**	**Negative**
No source of bias considered	6	1	2	5
Double blind trials	0	0	0	0
Observer blind trials	2	1	0	3
Non blind trials	4	0	2	2
Reporting quality 3 or more	0	1	0	1
Reporting quality 2 or less	6	0	2	4
Validity score 9 or more	1	1	0	2
Validity score 8 or less	5	0	2	3
European studies	2	1	1	2
Far east studies	4	0	1	3

Reporting quality using 0-5 scale [Jadad et al, 1996]; Validity scoring using 0-16 scale [Smith et al, 2000]; Geographical definitions [Vickers et al; 1998]

higher proportion of positive results for poor quality trials than for those of higher quality. This was the case whether the authors conclusions or the reviewers conclusions were used (Table 2). This is in agreement with what is expected from other studies of bias [1,2,5,6,11,13].

An interesting point raised in this review is that of the definition of acupuncture. Traditional Chinese acupuncture is usually defined as therapy interfering or enhancing energy flow along meridians in the body. Electrical stimulation of muscles via acupuncture needles produced muscle contraction, described as "pronounced" in one study [21]. Three studies mention the production of muscle stimulation, but the regimen used in all was likely to produce it. Where rehabilitation is aimed at returning normal muscle function, it becomes questionable whether the treatment under test truly is acupuncture. Not one of the studies, for instance, used sham acupuncture as a control, but rather used conventional treatment. Electrical stimulation of muscle versus no stimulation was under test in six of the seven trials, not the interference of energy flows along meridians.

So what evidence is there that acupuncture is beneficial in stroke? The answer is simple - absolutely none that is in any way convincing, despite there being six reports claiming that it is beneficial. The bottom line is that it's use is wasteful, and may be dangerous. At the very least it directs resources and effort away from interventions, in stroke or elsewhere, that are known to work and benefit patients.

Acknowledgements

The study was supported from Pain Research Funds.

References

1 Schulz KF, Chalmers I, Hayes RJ, Altman DG. **Empirical evidence of bias: Dimensions of methodological quality associated with estimates of treatment effects in controlled trials**. *JAMA* 1995, **273**: 408-12.

2 Carroll D, Tramèr M, McQuay H, Nye B, Moore A. **Randomization is important in studies with pain outcomes: systematic review of transcutaneous electrical nerve stimulation in acute postoperative pain.** Br J Anaesth 1996, **77**: 798-803.

3 Ernst , White AR. **Acupuncture for back pain: A meta-analysis of randomised controlled trials**. Arch Int Med 1998, **158**: 2235-2241.

4 Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJM, Gavaghan DJ, McQuay HJ. **Assessing the quality of reports of randomized clinical trials: is blinding necessary?** *Control Clin Trial* 1996, **17**: 1-12.

5 Khan KS, Daya S, Jadad AR. **The importance of quality of primary studies in producing unbiased systematic reviews**. *Arch Intern Med* 1996,**156** :661-6.

6 Moher D, Pham B, Jones A, et al. **Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses?** *Lancet* 1998, **352** :609-613.

7 Moore RA, Collins SL, McQuay HJ. **Variation in the placebo response rate: implications for individual trials and consequences for meta-analysis**. In *Proceedings of the 8th World Congress on Pain*, Progress in Pain Research & Management, Vol 8. Editors Jensen TS, Turner JA & Wiesenfeld-Hallin Z, IASP Press, Seattle, 1997.

8 Moore RA, Carroll D, Wiffen PJ, Tramèr M, McQuay HJ. **Quantitative systematic review of topically-applied non-steroidal anti-inflammatory drugs**. *BMJ* 1998, **316**: 333-8.

9 Moore RA, Gavaghan D, Tramèr MR, Collins SL, McQuay HJ. **Size is everything - large amounts of information are needed to overcome random effects in estimating direction and magnitude of treatment effects**. *Pain* 1998, **78**: 217-220.

10 Tramèr M, Reynolds DJM, Moore RA, McQuay HJ. **Effect of covert duplicate publication on meta-analysis; a case study**. *BMJ* 1997, **315**: 635-40.

11 Vickers A, Goyal N, Harland R, Rees R. **Do certain countries produce only positive results? A systematic review of controlled trials.** Control Clin Trial 1998, **19**: 159-166.

12 McQuay HJ, Moore RA. **An evidence-based resource for pain relief**. Oxford University Press March 1998.

13 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.

14 Jadad AR, McQuay HJ. **Meta-analysis to evaluate analgesic interventions: a systematic qualitative review of the literature**. *J Clin Epidemiol* 1996, **49**:235-243.

15 Smith L, Oldman A. **Acupuncture and dental pain**. Br Dent J 1999, **186**: 158-159.

16 Ernst E, White AR. **Acupuncture as an adjuvant therapy in stroke rehabilitation?** *Wien Med Wochenschr* 1996, **146**:556-8.

17 Li D, Wenduo L, Lihua W, Yulin Z, Huaizhong L. **Clinical observation on acupuncture therapy for cerebral hemorrhage.** *J Trad Chin Med* 1989, **9**:9-13.

18 Hu H, Chung C, Liu TJ, Chen RC, Chen C, Chou P, Huang W, Lin JCT, Tsuei JJ. **A randomized controlled trial on the treatment for acute partial ischaemic stroke with acupuncture.** *Neuroepidemiol* 1993, **12**:106-13.

19 Johansson K, Lindgren I, Widner H, Wiklund I, Johansson BB. **Can sensory stimulation improve the functional outcome in stroke patients?** *Neurology* 1993, **43:**2189-92.

20 Kjendahl A, Sällström S, Østern PE, Stanghelle JK, Borchgrevik CF. **A one year follow-up study on the effects of acupuncture in the treatment of stroke patients in the subacute stage: a randomized, controlled study.** *Clin Rehabil* 1997, **11**:192-200.

21 Gosman-Hedström G, Claesson L, Klingenstierna U, Carlsson J, Olausson B, Frizell M, Fagerberg B, Blomstrand C. **Effects of acupuncture treatment on daily life activities and quality of life: A controlled, prospective, and randomized study of acute stroke patients.** *Stroke* 1998, **29**:2100-8.

22 Si Q, Wu G, Cao X. **Effects of elecroacupuncture on acute cerebral infarction.** *Acupunct & Electro-*

therapeut Res Int J 1998, **23**:117-124.

23 Wong AMK, Su T, Tang F, Cheng P, Liaw M. **Clinical trial of electrical acupuncture on hemiplegic stroke patients.** *Am J Phys Med Rehabil* 1999, **78**:117-122.

24 Magnusson M, Johnasson K, Johansson BB. **Sensory stimulation promotes mornalization of postural control after stroke**. *Stroke* 1994, **25**:1176-80.

25 Shun Y, Yu, Z, Shun S, Yan C, Dai T, Liang L. **Observing the effect of the nail microcirculation of the hemiplagia patients by needling Tongtian and anterior Shencong to Xuanli.** *Acupunct Res* 1988, **13**:90-93.

26 Sällström S, Kjendahl A, Østern PE, Stanghelle JK, Borchgrevik CF. **Akupunkturbehandling ved hjernselag I subakutt fase/ En randomisert kontrollert studie.** *Tissskr Nor Laegeforen* 1995, **115**:2884-7.

27 Sällström S, Kjendahl A, Østern PE, Stanghelle JK, Borchgrevik CF. **Acupuncture in the treatment of stroke patients in the subacute stage: a randomized, controlled study.** *Phys Ther* 1996, **4**:193-7.

28 Deeks JJ, Glanville JM, Sheldon TA, Watt I. Undertaking Systematic Reviews of Research on Effectiveness: CRD Guidelines for Those Carrying Out or Commissioning Reviews. CRD Report 4. 1966 York: University of York.

Table 1: Details of studies, design and outcomes is shown overleaf.

Paper	Study entry criteria	Treatment groups	Design and outcomes
Li et al, 1989	Inclusion: Diagnosis of cerebral haemorrhage made by consideration of the course of illness, typical signs and symptoms supplemented by bloody CSF. Ill-defined patients were diagnosed with CT and those for whom CT or CSF diagnosis was impossible then diagnosis was reached by discussion of consultants. Exclusions not mentioned Age range: 26-85 Number: 92	Treatment group (N=46): Principal acupoints were Fengfu and Yamen which were needled once daily from admission with needles 12-16% of diameter of the neck. Control group (N=46): No neck acupuncture Both groups: Received acupuncture in 6-10 adjuvant points according to their specific conditions (e.g. hypertension). These sites were electrically stimulated for 10-20 mins once the patient's condition stabilised. One course of therapy lasted for 14 days and there followed a 3 day respite until the next course started (3-4 courses on average)	Randomised but method not stated Not blind Withdrawals reported - 22 patients died; 8 in the treatment group and 14 in the control group. Deaths were included in the analysis. No primary outcome stated Statistical methods not given Conducted in China
Hu et al, 1993	Inclusion: First stroke onset within 36 hrs Exclusions: 1) previous history of stroke; 2) CT showing cerebral haemorrhage or intercranial lesion other than stroke; 3) ischaemic stroke other than in middle cerebral artery territory; 4) TIA; 5) coma; 6) any other life-threatening illness; 7) significant systemic diseases or diseases which interfere with the assessment of stroke; 8) stroke without limb weakness Age range: 46-74 Number: 30	Treatment group (N=15): 7 body sites and 3 scalp points on the paralytic side were used routinely with electric stimulation at 9.4 Hz being applied to the scalp for 10 min and the body for 20 min during each session. No session exceeded more than 60 min and additional acupoints were only selected to meet certain requirements. Every other day for 4 weeks Control group (N=15): Same supportive treatment as treatment group, such as fluid therapy, prevention of complications and a standard rehabilitation program.	Randomised but method not stated Not blind Withdrawals - no drop outs or deaths occurred No primary outcome stated Statistical methods explicit Conducted in Taiwan
Johansson et al, 1993	Inclusion: Patients with hemiparesis within 10 days of stroke onset who; 1) could co-operate during the examination and tests; 2) had a paresis to the extent that they required help to eat, dress and walk. Exclusion: 1) Patients who could not manage ADL; 2) patients with pacemakers. Median age: 76 Number: 78 (51 with CAT scan)	Treatment group (N=38): Sites on both sides of the body were used with 10 acupoints used in total for 30 min per session with manual stimulation being used on all sites and electric stimulation at 2-5 Hz given to the two sites on the paralytic side. Intensity sufficient to produce muscle contraction. Treatment was given twice a week for 10 weeks. Control group (N=40): standard therapy alone. Both groups: Standard individual stroke rehabilitation including daily physiotherapy and occupational therapy.	Randomised but method not stated Not blind Withdrawals reported - by the end of the study (12 months post-stroke) 18 patients had died: 10 - cardiac disease; 2 - recurrent stroke; 1 each - pulmonary oedema, pneumonia, intestinal gangrene, intestinal haemorrhage. Deaths were not included in analysis. No primary outcome stated Statistical analysis descrbed explicitly Conducted in Sweden
Kjendahl et al, 1997	Inclusion: Patients were enrolled on attending the hospital for rehabilitation following their first ever stroke. Interval from stroke onset was 15-71 days. Exclusion: Subarachnoid bleeding or other significant diseases. Age range: 35-72 Number: 41	Treatment group (N=24): body and scalp at traditional Chinese points. Moxibustion or electrical stimulation at 2-4 Hz in some patients, 3-4 times per week for 30 min for 6 weeks. Control group (N=21): Patients were given an individually designed rehabilitation programme in the same way as treated patients.	Randomised but method not stated Observer blind for motor assessment score only Withdrawals reported - 4 drop-outs; 2 per group. 2 patients died; 1 from pneumonia and 1 from MI. 2 patients withdrew; 1 disliked needles and 1 could not be retested due to early discharge from hospital. Drop-outs were not included in the analysis. No primary outcome stated Statistical methods explicit Conducted in Norway

Paper	Outcome Measures	Results	Adverse Effects	Quality Score	OPVS	Conclusion	
						Author	Reviewer
Li et al, 1989	Patient assessment: Scoring of Disease Condition in Cerebro-Vascular Accidents (devised by authors). Consists of scoring for consciousness, speech, motor function of the tongue and facial muscles and the function of upper and lower extremities (total of 100 points with that score as fully normal). Grading: Ineffective - increase <15; Effective - increase 15-29; Markedly effective - increase >30; Essentially cured - 90-97; Cured >97 Patients were assessed at the start of the trial, at the end of each course of treatment and at discharge.	Overall, cured/essentially cured; treatment group - 50%, control - 19.6%: p<0.01 Effects for disturbances of consciousness, completely recovered: treatment s.d. p<0.01 Effects for aphasia: treatment s.d. p<0.01 Effects for paralysis, muscle strength > 3 (0-5, 5=normal), upper and lower limbs: treatment s.d. p<0.01	None reported	2	8	Positive	Positive
Hu et al, 1993	Neurological deficit: Scored using the system developed and used in the Multi-Centre Trial of Haemodilution in Acute Ischaemic Stroke by the Scandinavian Stroke Group. Assessed at days 1, 2, 3, 7, 14, 21, 28, 90 Functional impairment: Barthel index score. Assessed at days 7, 14, 21, 28, 90	Neurological,median difference in improvement, 28 days; acu s.d. p=0.02 Neurological, median difference in improvement, 90 days; acu s.d. p=0.009 Functional, change in scores, 28 and 90 days; no s.d.	None reported	2	4	Positive	Negative: positive result depended only on sub group analysis and on only some outcomes
Johansson et al, 1993	Mobility score - included sitting/standing up, turning in bed. Walking (0=inability to walk, 6=normal). Balance (21 max score). Motor function - evaluated 31 active different movements of the arms, wrists, hands and legs and 4 rapid agonist-antagonist movements (max 105=normal function). All tested at baseline, 1month and at 3 months. Physical disability: Barthel's index. Assessed at baseline, 1, 3 and 12 months. Quality of life: Nottingham Health Profile. Contains questions on energy, physical mobility, sleep, pain, emotional reactions and social isolation (Scored 0-100 with higher scores showing more problems). Assessed at 3, 6 and 12 months.	Motor function, 1 month; acu -65.2 (5.0): control - 49.8 (6.0): nsd; 3 month; acu - 73.8 (4.9): control - 62.1(5.8): nsd Walking, 1 month; acu - 3.2 (0.3): control - 22 (0.3): p<0.01; 3month; acu - 3.9 (0.3): control - 2.9 (0.3): p<0.004 Balance, 1 month; acu - 14.6 (0.8): control - 11.0 (0.8): p<0.001; 3 month; acu - 16.1 (0.7): control - 12.9 (0.7): p<0.001 ADL, 1 month; acu - 69.4 (3.0): control - 60.6 (3.4): p<0.05; 3 month; acu - 90.4 (2.2): control - 72.4 (3.2): p<0.0001; 12 month; acu - 92.0 (2.9): control - 71.3 (4.0): p<0.0001 NHP - at 3 month, energy, mobility, emotion and social isolation significantly different in acu group. At 6 month, energy, mobility, emotion, social isolation and sleep significantly different. At 12 month, mobility and emotion significantly different.	None reported	2	8	Positive	Positive
Kjendahl et al, 1997	Motor function: Motor Assessment Scale - assesses turning in bed, sitting, standing up, walking, balance in sitting, activities of the upper arm, wrist and hand. (0-48, 48=normal) ADL: Sunnaas index of ADL - includes, eating, continence, indoor mobility, toilet management, transfer, dressing and undressing, hygiene, bath/shower, cooking, housework, outdoor mobility and communication. (0-36, 36=normal) Quality of life: Nottingham Health Profile. Contains questions on energy, physical mobility, sleep, pain, emotional reactions and social isolation (Scored 0-100 with higher scores showing more problems). Assessed at baseline, 6weeks and 12 months.	MAS, 6 wk; acu - 29.1 (9.7): control - 26.3 (11.1): p=0.002; 12 months; acu - 35.0 (10.0): control - 29.6 (11.4): p=0.001 ADL, 6 wk; acu - 24.8 (6.6): control - 24.3 (6.3) nsd; 12 months; acu - 31.9 (4.9): control - 26.5 (5.8): p=0.0002 NHP, 6 wk; acu - 11.4 (11.4): control - 21.6 (20.1): p-0.009;12 months; acu - 5.91 (6.8): control - 21.0 (17.6): p<0.0001	None reported	2	7	Positive	Negative. Statistical significance only obtained on small differences between start and finish points.

Paper	Study entry criteria	Treatment groups	Design and outcomes
Gosman-Hedstrom et al, 1998	Inclusion: >39 years old with an acute focal ischaemic non-haemorrhagic lesion <1wk before randomisation. Paresis to an extent that the patient required help to walk, eat and dress and the patient had to be able to co-operate mentally and be willing to take part in the study. Exclusion: 1) other severe disease requiring hospital/nursing home care; 2) severe aphasia or unconsciousness; previous cerebral lesion; 4) pacemaker Mean age: 77 years Number: 104	Deep acupuncture group (N=34): Both sides of the body were used with 10 points in total. The nonparalytic needles were stimulated manually ever 5mins and those on the paralytic side were electrically stimulated at 2Hz tp produce pronounced muscle contractions. Each treatment was given for 30mins. Superficial acupuncture (N=33): One needle was inserted just under the skin on each limb and left for 30 min with no stimulation. Acupuncture treatment was given twice a week for 10 weeks in both groups. Conventional stroke rehabilitation group (N=37). Conventional rehabilitation was used for all three groups	Randomisation by computer Observer blind for all outcomes except Nottingham Health Profile Withdrawals reported - 4 withdrawals; 3 in the deep acupuncture group (2 - disliked treatment; 1 - developed arm infection); 1 in superficial acupuncture group wished to stop. All withdrawals were included according to the intention-to-treat principle. 19 patients died in total; 4 in the deep acupuncture group, 10 in the superficial group and 5 in the no treatment group. All patients bar one were completely assessed. Deaths were as follows; 11 - cardiac death; 2 - pneumonia; 4 - cerebral infarction; 1 - renal insufficiency; 1 - GI bleeding No primary outcome stated Statistical methods explicit Conducted in Sweden
Si, Wu and Cao, 1998	Inclusion: 1) confirmation of cerebral infarction by CT; 2) with in 7 days of onset; 3) no apoplectic coma; 4) hemiplegia (muscle strength >3 where 0=normal and 6=inability to move) Exclusions: none given Age range: 58-78 years Number: 42	Treatment group (N=20): Acupuncture on 6 sites on the paralytic side, manually stimulated then stimulated at 5/45 Hz for 30 min. Once per day for 5 days then rested for 2 days repeated until discharge. Control group (N=22): Treated with heparin, low molecular weight dextran and nimodipine, as with acupuncture Stay in hospital 23-49 days.	Randomised but method not stated Observer blind outcome assessment Withdrawals not discussed No primary outcome stated Statistical methods not described Conducted in China
Wong et al, 1999	Inclusion: 1) CT evidence of stroke; 2) hemiplegia; 3) first stroke; 4) physically stable 10-14 days after onset; 5) clear consciousness and no complications during the medical course. Exclusions: none given Age range: 21-80 years Number: 118	Treatment group (N=59): Adhesive surface electrodes placed over 8 sites on the paralytic side (4 per limb) and stimulated at 20-25 Hz for 30mins, sufficinet to induce muscle contraction. 5 times per week for 2 weeks. Control group (N=59): Same supportive treatment and a standard rehabilitation programme including active physiotherapy and occupational therapy every day until discharge from hospital.	Randomised but method not stated Not blind Withdrawals not discussed No primary outcome stated Statistical methods explicit Conducted in Taiwan

Paper	Outcome Measures	Results	Adverse Effects	Quality Score	OPVS	Conclusion	
						Author	Reviewer
Gosman-Hedstrom et al, 1998	Functional outcome: Based on the neurological score according to the Scandinavian Stroke Group - includes, motor function of arm, hand and leg, ambulation, orientation and speech (max 48 points) ADL outcome: Barthel index (assesses 10 activities; max 100 pts) and Sunaas index of ADL (12 activities, max 36 pts) with higher scores indicating less problems. Quality of life: Nottingham Health Profile. Contains questions on energy, physical mobility, sleep, pain, emotional reactions and social isolation (Scored 0-100 with higher scores showing more problems). Evaluated - 3 days after randomisation, 3 weeks, 3 and 12 mths.	Because superficial acupuncture and conventional treatment groups showed no significant difference, they were combined for comparison with deep acupuncture group. Neurological, 3 months; deep - 9.27 (9.23): superficial - 9.17 (7.37): no - 11.07 (6.88);12 months; deep - 11.23 (9.64): superficial - 9.96 (8.67): no - 11.75 (7.48) Barthel, 3 months; deep - 38.18 (24.77): superficial - 32.0 (27.34): no - 40.17 (20.02); 12 months; deep - 41.94 (25.78): superficial - 37.17 (26.28): no - 42.86 (22.67) Sunnaas, 3 months; deep - 12.21 (7.27): superficial - 10.23 (8.33): no - 13.37 (6.66); 12 months; deep - 13.55 (7.74): superficial - 12.30 (8.36): no - 14.89 (7.99) NHP - only significant difference is no acu at 12 months for physical disability which is significantly worse ($p<0.05$)	None reported	3	11	Negative	Negative
Si, Wu and Cao, 1998	Chinese Stroke Scale (0-43, 0=normal function) Assessed - admission and after treatment finished (23-49 days after admission).	CSS, final results; acu - 15.8 (6.4): control - 19.9 (9.1) CSS, score changes; acu - 8.2 (3.4): control - 5.1 (3.4): $p<0.01$ Motor shoulder, hand and leg improved in acu group ($p<0.05$)	None reported	1	10	Positive	Negative. There was no difference between patients at beginning and end of study.
Wong et al, 1999	Neurological Status: Brunnstrom's stages from I to IV as a score of 1 to 6 (6=normal) Functional assessment: Functional independence measure - assesses self-care, sphincter control, mobility, locomotion, communication and social cognition (score from low of 18 to high of 126) Evaluated - admission and discharge (aprox 21-40 days)	Neurological - Upper limb, discharge; acu group - 3.1 (1.1): control - 2.6 (0.9): $p=0.02$; Lower limb, discharge; acu - 3.9 (1.3): control - 3.3 (1.0): $p=0.004$ Functional - Discharge; acu - 69.6 (21.0): control - 61.8 (16.6): $p=0.02$	None reported	1	8	Positive	Negative. Statistical analysis and reported results are not believable. For instance Table 4 has exactly the same mean symptom improvement score for seven different items