Modellvorhaben Akupunktur – a summary of the ART, ARC and GERAC trials

Mike Cummings

In October 2000 the German Federal Committee of Physicians and Health Insurers recommended that special Model Projects on Acupuncture (”Modellvorhaben Akupunktur”) be developed in order to determine the evidence-based role of acupuncture in the treatment of certain illnesses. This paper presents a summary of the main randomised controlled trials performed as part of these projects, and the associated economic analyses.

Overall the results show that acupuncture is effective in practice for a range of chronic conditions, and it seems likely to have acceptable cost utility (at least at a rate of €35 per session). Sham acupuncture, in the form of minimal off-point needling in a therapeutic context, also appears to be effective, being no different to prophylactic medication in migraine, and superior to guideline-based standard care in chronic low back pain.

In conclusion, acupuncture appears to be effective in a range of chronic conditions and it seems to have acceptable cost-effectiveness in Western health economic terms. These programmes of research do not confirm the hypothesis that needling at specific points is essential to achieve satisfactory clinical effects of acupuncture. Sham acupuncture, in the form of minimal off-point needling in a therapeutic context, is unlikely to be an inactive placebo.

In April 2006, the German health authorities decided that acupuncture would be included into routine reimbursement by social health insurance funds for chronic low back pain and chronic osteoarthritis of the knee.

Three large research programmes investigating the efficacy, effectiveness, cost-effectiveness and safety of acupuncture treatment for certain chronic conditions (“Modellvorhaben Akupunktur”) have been conducted in Germany since October 2000.1 These programmes were initiated after the German Federal Committee of Physicians and Health Insurers determined, in October 2000, that the scientific evidence supporting the use of acupuncture was not sufficient to justify routine reimbursement within the German healthcare system. Formerly, some of the cost of acupuncture treatment was covered by the German statutory health insurance funds, provided that the acupuncture was performed by physicians with at least 140 h of acupuncture training.

Following the decision of the German Federal Committee of Physicians and Health Insurers in October 2000, reimbursement for acupuncture treatment was only possible for patients suffering from certain chronic conditions (knee osteoarthritis, low back pain, migraine, tension-type headache), and only if the physician performing the acupuncture participated in one of the three research programmes. The results discussed in this paper are from the controlled trials that made up the core of these research programmes: the Acupuncture Randomised Trials (ART),2–5 the Acupuncture in Routine Care (ARC) studies,6–14 one comparative trial (COMP),15 and the German Acupuncture trials (GERAC).16–19

These Modellvorhaben (trial phases) were funded by a number of the German statutory health insurance funds, and they were organised by groups of researchers and physicians based at three large German universities: the Charité University, Berlin (ART and ARC); the Technical University, Munich (ART and COMP); and the University of Bochum (GERAC).

METHODS

ART

These were four randomised controlled trials (RCTs) with roughly 300 subjects in each. They were performed principally as efficacy trials in four conditions: migraine6; tension-type headache6; chronic low back pain6; and osteoarthritis of the knee. One of each of the trials followed the same design: three parallel arms with a 2:1:1 distribution of subjects, so that there were approximately 150 subjects in the real (verum) acupuncture arm, and 75 in the others — the minimal acupuncture and waiting list arms. The acupuncture involved deep needling to classical acupuncture points with manipulation of the needles to produce de qi — a characteristic needling sensation. Twelve treatments were given over 8 weeks. Minimal acupuncture involved superficial needling to standardised sites that were not near to any recognised acupuncture points. The waiting-list group received acupuncture 2 or 3 months after randomisation, that is, after the data were collected for the primary outcome.

The primary outcomes were short term, just after the interventions at around 8 weeks, although outcomes were also assessed at 26 and 52 weeks from baseline.

The primary outcome for ART migraine was the difference in number of days with headache of moderate or severe intensity between the 4 weeks before randomisation (baseline phase) and weeks 9–12 after randomisation. Responders were defined (post hoc) as those with a 50% reduction or greater in days with moderate or severe pain (headache). The primary outcome and responder rates for ART tension-type headache were the same, with an additional comment that patients with missing data were automatically counted as non-responders.

The primary outcome in ART low back pain was the change in low back pain intensity from baseline to the end of week 8 after randomisation, as measured by a visual analogue scale (range, 0–100 mm), and responders were defined (post hoc) by at least 50% reduction in pain intensity. Finally, the primary outcome measure in ART knee osteoarthritis was the change in the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) between baseline and week 8 after randomisation, and responders were defined (post hoc) by a decrease of at least 50% in their WOMAC index score.

The ART programme trials were performed across between 18 and 30 outpatient centres across Germany: ART migraine 18; ART tension-type headache 28; ART low back pain 30; and ART knee osteoarthritis 28.

COMP

This was a single comparative trial of acupuncture and metoprolol (100–200 mg) in migraine prophylaxis.20 It was designed as an equivalence trial, with a recruitment target of 480. The acupuncture treatment consisted of at least eight to a maximum of 15 sessions of 20 to 30 minutes’ duration, administered over a period of 12 weeks. At least six needles were used per session, with manual stimulation at least once to achieve de qi. The metoprolol intervention was
based on the recommendations of the German Migraine and Headache Society and consisted of metoprolol 100–200 mg daily for 12 weeks.

The primary outcome measure was the difference in number of days with migraine between the 4 weeks before randomisation (baseline) and weeks 9–12 after randomisation as reported by the patient in a standardised headache diary. Responders were defined as those with at least a 50% reduction in migraine attacks.

GERAC

There were four GERAC trials with up to 1000 subjects in each. They were designed as comparative trials with three equal parallel arms: acupuncture versus sham acupuncture versus standard care (note that the terminology “sham” is used rather than “minimal”). They were performed in migraine, tension-type headache, chronic low back pain, and knee osteoarthritis.

Rather like the ART trials, the real acupuncture involved deep needling to classical acupuncture points with manipulation of the needles to produce *de qi*. Ten treatments were given over 6 weeks, with the option to extend treatment by a further five sessions for partial response. Sham acupuncture involved superficial needling to standardised sites that were not near to any recognised acupuncture points. The standard care arms used best conventional care based on guidelines where available: GERAC migraine — beta blockers first choice, flunarizine second, valproic acid third; GERAC tension-type headache — the intention was to use amitriptyline; GERAC chronic low back pain — multimodal treatment programme including physiotherapy, exercise and non-steroidal anti-inflammatory drugs (NSAIDs); GERAC knee osteoarthritis — physiotherapy, physician visits, NSAIDs (in this trial all groups had six sessions of physiotherapy, and acupuncture groups were allowed limited NSAIDs as rescue medication).

The primary outcomes were measured around 6 months from baseline, although secondary outcomes were measured at 6 weeks and 3 months as well.

The primary outcome in GERAC migraine was the difference in migraine days between 4 weeks before randomisation and weeks 23–26 after randomisation, and response was defined as a reduction in the number of migraine days by 50% or more. GERAC tension-type headache was somewhat different in that the response (defined as >50% reduction in number of headache days per 4 weeks from baseline to 6 months) was the primary outcome, and all minor variations from protocol resulted in patients being classified as non-responders.

In GERAC low back pain the primary outcome was response after 6 months, defined as 33% improvement or better on three pain-related items on the Von Korff Chronic Pain Grade Scale questionnaire (CPGS) or 12% improvement or better on the back-specific Hanover Functional Ability Questionnaire (HFAQ). Patients who were unblinded or who used (disallowed) co-interventions during follow-up were classified as non-responders regardless of symptom improvement.

In GERAC knee osteoarthritis the effect on pain and function was measured with the WOMAC score (total score and the subscales were standardised to 0–10).

“Success” rates were calculated according to a change of at least 36% from baseline WOMAC scores at 13 and 26 weeks after the start of treatment. Patients with missing data were considered to have had treatment failure.

The GERAC trials were performed across between 122 and 340 practices across Germany: GERAC migraine 149; GERAC tension-type headache 122; GERAC chronic low back pain 340; GERAC knee osteoarthritis 515.

ARC

The ARC studies were a series of large to very large pragmatic RCTs, with associated non-randomised cohorts. They used a standard design and included detailed economic analysis from a societal perspective. Subjects insured by one of the participating social health insurance funds were recruited by general practitioners across Germany for acupuncture treatment of either: osteoarthritis of the hip or knee; chronic neck pain; chronic low back pain; chronic headache; dysmenorrhoea; allergic rhinitis; or asthma (awaiting publication). If they agreed to be randomised, they either received 15 sessions of manual acupuncture over 3 months, or they waited 3 months for acupuncture treatment. If they expressed a strong preference for acupuncture and declined to be randomised, they received acupuncture treatment immediately. There was no standardisation of treatment, but only manual acupuncture was allowed.

Outcomes were measured at 3 and 6 months. After 3 months the group randomised to usual care alone were given acupuncture treatment. The primary outcomes were all set at 3 months. ARC chronic headache used the reduction in days with headache per month. ARC low back pain measured back function assessed by the HFAQ. ARC osteoarthritis used the change in WOMAC score, and ARC chronic neck pain used a validated neck pain and disability scale (NPAD). In ARC dysmenorrhoea the main outcome

![Figure 1](image-url) **Figure 1** Responder rates in the Acupuncture Randomised Trials (ART) trials after 8 weeks from baseline (9–12 weeks in ART migraine and tension-type (TT) headache); responder rates were defined (post hoc) as a 50% or greater reduction in the primary outcome measure. Acupuncture and minimal acupuncture were significantly superior to waiting list in all trials. Acupuncture was superior to minimal acupuncture only in ART knee osteoarthritis (OA).
was the average pain intensity during the last menstruation before assessment measured on a numeric rating scale. ARC allergic rhinitis used the Rhinitis Quality of Life Questionnaire (RQLQ).

As part of the ARC programme of studies, additional measurements were performed to assess quality of life (QoL), costs and the cost-effectiveness relationship of routine care plus acupuncture compared with routine care alone. QoL was assessed with the Short Form (SF)-36 questionnaire, using the subscales and the components scales. The SF-36 also served as the basic benefit estimator for the cost-effectiveness analyses. At baseline and at 3 months the patients completed questionnaires which assessed the QoL over the previous 7 days. The costs considered were measured in societal perspective and included the direct healthcare-related costs of acupuncture (cost of each acupuncture session was €35), physician visits and hospital stays, and any drugs prescribed. In addition to health insurance costs, the indirect costs caused by lost workdays were also taken into account. These were estimated to be approximately €78 per lost workday. Additional analyses were performed to estimate cost utility in the case of higher costs and better medical outcome. QoL measures using SF-36 were converted to quality-adjusted life-years (QALYs), and the excess cost in the acupuncture group in each study was divided by the increment in QALYs gained in the acupuncture group compared with the usual care group. This gave an incremental cost-effectiveness ratio (ICER) expressed as a cost per additional QALY.

**RESULTS**

**ART**

An overview of the main results of the ART trials is shown in fig 1, expressed as responder rates for comparison across the different conditions. In all four trials there were significant short-term differences between acupuncture and waiting list, but there was a significant difference between acupuncture and minimal acupuncture only in ART knee osteoarthritis. There were significant short-term differences between minimal acupuncture and waiting list in all four trials. Treatment effects were maintained in the acupuncture and minimal acupuncture groups at long-term follow-up (21–24 weeks in ART migraine and tension-type headache; 52 weeks in ART low back pain and knee osteoarthritis).

**Figure 2** Responder rates in the comparative trial (COMP) at 9–12 weeks from baseline, and the German Acupuncture trials (GERAC) at 6 months from baseline; responder rates were defined as: ≥50% reduction in migraine days (COMP and GERAC migraine); ≥50% reduction in headache days (GERAC tension-type (TT) headache); ≥33% improvement on Chronic Pain Grade Scale questionnaire (CPGS) or ≥12% improvement on the Hanover Functional Ability Questionnaire (GERAC low back pain); and ≥36% improvement in Western Ontario and McMaster Universities Osteoarthritis Index (GERAC knee osteoarthritis (OA)). Acupuncture and sham acupuncture were both significantly superior to standard therapy in GERAC low back pain and GERAC knee OA. There were no other statistically significant differences between groups.

**Figure 3** Percentage improvement in the primary outcome measure at 3 months from baseline in the Acupuncture in Routine Care (ARC) trials ("control after acupuncture" is 6 months from baseline — 3 months usual care followed by 3 months acupuncture treatment). Numbers in brackets are those randomised (r) followed by the total sample including the non-randomised cohort. In all six trials there was a very highly significant difference between acupuncture and usual care alone at 3 months (p<0.001).
Rates were defined as: expressed as responder rates. Responder rates were 53.1% for acupuncture, 44.2% (minimal acupuncture), and 27.4% (conventional therapy). The differences in terms of the response to acupuncture when compared with the standard therapy group were superior to minimal acupuncture for most secondary outcomes, including headache days (1.8 fewer; p = 0.004) and the International Headache Society response criterion (>50% reduction in headache days: 66% vs 55%, risk difference 12%; p = 0.024).

In GERAC low back pain the response rates at 6 months were 47.6% (acupuncture), 44.2% (minimal acupuncture), and 27.4% (conventional therapy). The differences among groups were as follows: acupuncture versus minimal acupuncture, 3.4% (p = 0.59); acupuncture versus conventional therapy, 20.2% (p<0.001); and minimal acupuncture versus conventional therapy, 16.8% (p<0.001).

In GERAC knee osteoarthritis the success rates were 53.1% for acupuncture, 51.0% for minimal acupuncture, and 29.1% for conservative therapy. Acupuncture groups had higher success rates than the conservative therapy group (relative risk (RR) for acupuncture compared with conservative therapy, 1.75 (p<0.001); RR for sham acupuncture compared with conservative therapy, 1.75 (p<0.001)). There was no difference between acupuncture and minimal acupuncture (RR, 1.01 (p = 0.48)).

**DISCUSSION**

The results show that acupuncture is effective in practice for a range of chronic conditions, and it seems likely to have acceptable cost utility (at least at a rate of €55 per session). Sham acupuncture, in the form of minimal off-point needling in a therapeutic context, also appears to be rather effective, being no different to prophylactic medication in migraine (GERAC migraine), and being superior to guideline-based standard care in chronic low back pain (GERAC low back pain).

**Education**

**Highlights**

- Acupuncture was more effective than placebo for all conditions.
- Acupuncture was superior to guideline-based standard care in chronic low back pain.
- Acupuncture was effective in migraine, tension-type headache, knee osteoarthritis, and low back pain.
- The results support the use of acupuncture for a range of chronic conditions.
Summary points

- The Modellvorbahen Akupunktur are three large research programmes that have attempted to investigate the efficacy, effectiveness, cost-effectiveness and safety of acupuncture treatment for certain chronic conditions including headache, back pain, knee pain and neck pain.
- The results demonstrate that acupuncture is effective in practice.
- Acupuncture (at a rate of €35 per session) in addition to usual care seems likely to have acceptable cost utility when compared with usual care alone in these conditions.
- Sham acupuncture, in the form of minimal off-point needling in a therapeutic context, appears to be no different to prophylactic medication in migraine, and superior to guideline-based standard care in chronic low back pain, hence it is unlikely to be an inactive placebo.
- In patients recruited to acupuncture trials, the response to treatment does not differ between those that agree to be randomised and those that do not.
- Length of training in acupuncture does not seem to influence the results of treatment.

Mike Cummings

Correspondence to: Mike Cummings, BMAS, 60 Great Ormond Street, London WC1N 3HR, UK; BMASLondon@aol.com

Acknowledgements: Thanks to Claudia Witt and Benno Brinkhaus for providing data where necessary to complete the presentation of results.

Competing interests: MC is medical director of the BMAS. This role involves running short training courses for regulated healthcare professionals in Western medical acupuncture.

References