Practice Reports

EVIDENCE-BASED ACUPUNCTURE? A CHALLENGE AHEAD

HUGH MACPHERSON

Abstract

Acupuncture's increasingly widespread use has been paralleled by calls for more rigorous evidence of its effectiveness. In this article, some of the challenges of developing a strong evidence base are explored. As acupuncture is a complex intervention, a case is made for the appropriateness and relevance of pragmatic rather than experimental randomised controlled trials in the evaluation of acupuncture. Within this context there remain key challenges associated with ensuring treatment within a trial is sufficiently flexible, so that it can match the expected variability in patients, that outcomes measures capture the broader range of changes that can be experienced by patients, and that the therapeutic relationship is not compromised by the trial design. The central theme is the importance of protecting the integrity of the acupuncture, and the underlying principles upon which it is based.

In recent years there have been increasingly insistent calls for more compelling evidence of the effectiveness of acupuncture.¹ The most common question is 'Does it work?' And if it does, further questions follow, such as 'What is the mechanism?' 'Is it safe?' 'Is it cost effective?' This expectation of better evidence is fuelled by a number of underlying trends. First, the growing popularity of acupuncture, and its widespread use largely outside national health institutions in Europe and America, has created a sense that it is a public health issue. Second, prospective patients are willing to be responsible for their own health and therefore want to be better informed. Third, the medical establishment is interested in what techniques can be extracted and assimilated into conventional medicine, and is not particularly in favour of acupuncture being practised according to its traditional principles, especially by practitioners who are not first and

---

¹ National Institutes of Health 1998. World Health Organisation 2002.

foremost physicians. Fourth, there are a growing number of acupuncturists who, based on their immediate experience as practitioners, have a desire to 'prove' that acupuncture works. And lastly, there are policy and decision makers who are asking for evidence in order to decide whether to widen healthcare provision of acupuncture to specific groups of patients and/or for the treatment of specific conditions.

These expectations of better evidence of the effectiveness of acupuncture have in part been related to the emergence of the current biomedical fashion for 'evidence-based medicine'.² As originally conceived, the concept of evidence-based medicine was a sensible and practical approach to sifting and evaluating the clinical impact of treatments, thereby finding out what did, and did not, work. Clearly it is in the patients' interest to implement what works, whether the intervention is a new or existing one, and discontinue what does not.

The conventional medical literature abounds with treatments whose damaging effects were only discovered through a ruthless assessment of the evidence. A recent example of this has been the discovery that a policy of prescribing steroids to people with head injuries may have caused 10,000 unnecessary deaths.³ In the early days of evidence-based medicine, it was understood that a broad array of evidence should be incorporated in making any judgement of effectiveness. However, in more recent years an evidence hierarchy has emerged, where the only evidence considered of value is to be found from the results of randomised controlled trials. A whole industry of systematic reviews of such trials is associated with this trend. However too narrow an interpretation of what constitutes 'evidence' has created a climate whereby quantitative data from observational studies or qualitative data from interviews might be relegated to positions that might not even register on the evidenced-based map.

It is the purpose of this article to explore some of the challenges associated with developing a stronger evidence base for acupuncture. In this endeavour, the ideas I am putting forward are based on my experience of practice in the United Kingdom, North America and

---

² Sackett et al. 1996.

³ Sauerland and Maegele 2004.

Australia, and from this perspective I will focus on a number of key methodological challenges associated with developing the evidence base. I will identify some of the pitfalls and potentials of the randomised controlled trials in the clinical evaluation of acupuncture. And I will outline the need for a broader approach to evidence, in particular one that takes cognisance of the integrity of the acupuncture as it is practised in the West, and therefore incorporates important aspects of the experience of treatment as voiced by patients.

Complexity and diversity within acupuncture

Acupuncture is practised in a wide range of styles which come from many different traditions. New styles continue to evolve; for example 'auricular acupuncture' was only developed in the 1950s, and 'trigger point acupuncture' even more recently. While the practice of acupuncture is usually characterised by the insertion of needles at specific points to elicit improvement in health, there may be little similarity in theoretical underpinnings. Needles may be inserted briefly, or for a period of time; inserted to some depth or superficially, or even not at all; inserted at acknowledged acupuncture points or almost anywhere. In terms of the intention of the practitioner, the goal may be to move the *qi* and balance *yin* and *yang*, or it may be to release endorphins and boost neurotransmitters. This raises one of the first of the evidence challenges: which style of acupuncture should be investigated, and can this style be sufficiently well described so that the putative benefits can be reliably replicated.

Individual acupuncturists bring to their work a raft of prior experiences, both personal and professional. Their practice of acupuncture is known to incorporate many types of additional treatments provided as adjuncts to their primary healing art. Even when two acupuncturists have been trained within the same tradition, their treatments may well be different, and one individual practitioner might have a more positive effect on some patients than on others. Added to their individuality is often a passionate belief by practitioners that what they are doing is for the best, and doing anything different will diminish the therapeutic gain. That acupuncturists largely work outside state provision of health care in Europe and America means that they have also probably been outside the scrutiny of line

managers and peers.⁴ This has led to acupuncturists fostering an independence of spirit and a determined self-belief.

The patients who receive acupuncture are also individuals who have had their own unique experiences of ill health that have brought them to their acupuncturists. While for some patients, their condition may be a simple one, such as a sprained ankle, most acupuncture patients will have illness episodes of considerable duration and characterised by a range of symptoms, the constellation of which is interpreted as one or several overlapping patterns of disharmony or 'syndromes' in traditional Chinese medicine. This means that in normal practice no two patients are treated alike, and none can be expected to respond alike. This feature of practice is what makes the work of acupuncturists endlessly fascinating as each therapeutic encounter unfolds. And this unfolding often involves a journey for the patient, either a journey back to themselves, as in a typical statement, 'I've got my old self back again', or into their future as new possibilities open up.

Acupuncture, then, is a complex process with multiple interlocking dimensions. Any evaluative process must take into account these dimensions or something important at the heart of acupuncture may be lost. For example the evaluation of acupuncture needs to take into account the clinical context, so that the integrity of the acupuncture is maintained within the evaluation process. In addition, the patients who are being assessed in a clinical evaluation must have the same scope to interact with their acupuncture practitioner as they normally would. This means that they will bring to the therapeutic encounter their beliefs, whether sceptical or not, their attitudes, their relative willingness to engage with the healing process, and their ability to respond whether emotionally and physiologically. In so far as is possible with the above caveats, we would ideally want to develop evidence that is sufficiently robust to meet the current standards of evidence-based medicine. Therefore there is a need to think carefully about our approach in constructing an evaluation process. In addressing the challenges of evidence-based acupuncture, we also want to ensure that we are developing acupuncture-based evidence.

---

⁴ MacPherson et al. 2004a.

Pragmatic or experimental trials

For evidence to be convincing and compelling, the randomised controlled trial has much going for it. It is the methodology of choice when we want to unequivocally establish an association between cause (acupuncture) and effect (impact on health). The existence of a control group means that any benefits beyond those due to natural recovery can be measured, because we know that natural recovery has a significant impact on changes to health over time, so controlling for this is clearly important. However, the randomised controlled trial can be a disaster when applied without thought and sensitivity to the complex processes of acupuncture as described above.

What is less commonly known is that there are two basic types of randomised controlled trial.⁵ First there is the *experimental* trial which is designed to minimise all variables except the one which is being 'manipulated'. Ideally this trial would be like a laboratory experiment. By providing two interventions that differ only in terms of the 'active ingredient', it is possible to establish whether the difference in outcome can be ascribed to the active ingredient. To minimise the influence of all but the effect of the acupuncture needling *per se*, researchers attempt to standardise the components of the intervention. This is facilitated by blinding or masking the patients in the control group who only receive a 'placebo'-type sham intervention. This approach to research into acupuncture is based on the assumption that it is possible to deliver the 'specific' (needling) aspects of acupuncture separate from the other 'non-specific' aspects of the treatment. Blinding of practitioners is clearly impractical. The blinding of patients to which treatment they have been allocated and standardising their experience of their acupuncturist is widely promoted by medical researchers. However sham acupuncture with patient blinding necessarily creates an artificial and distorted treatment process which will tell us relatively little about the normal practice of acupuncture, as well as provide meaningless data for an analysis of cost-effectiveness.

The second type of randomised controlled trial is known as a *pragmatic* trial⁶ where the comparison is between an intervention such as

---

⁵ Roland and Torgerson 1998.

⁶ Garvey, Marks and Wiesel 1989.

acupuncture (ideally as it is provided in an everyday context) with some other intervention, such as standard conventional medical care. With this design it is possible to build into the trial design the necessary safeguards to protect the integrity of the acupuncture and the key components that contribute to the patient's experience.

Experimental trials tend to be more appropriate for well-defined and simple interventions that have a short acting effect, so that complicating factors are minimised. Pragmatic trials are better suited to evaluating complex interventions over a longer time period. Experimental trials have a place in exploring theoretical propositions in an artificial environment under very rigorous constraints, which means that they have the potential of providing highly rigorous evidence, but evidence that may not be so relevant to real world contexts. A classic example is the research into the acupuncture point Pericardium 6 where there is now 'high level' evidence of effectiveness when using this one point for post-operative nausea. Nevertheless, Pericardium 6 is seldom used as a single point for post-operative nausea in everyday practice. Pragmatic trials however afford a more appropriate approach to developing evidence of acupuncture's impact both clinically and economically, evidence that is relevant to the everyday clinical context which is characterised by complex and chronic conditions. If well designed, pragmatic trials can still carry sufficient rigour to contribute to the systematic reviews that are the current building blocks of the evidence base.

The contrast between these two approaches can be illustrated by reference to two randomised controlled trials of acupuncture, one for low back pain and one for headache. In an experimental trial conducted by Garvey, Marks and Wiesel, a single steroid injection for low back pain was compared to a single insertion of an acupuncture needle.⁷ Two weeks later, patients who received the acupuncture were found to be in significantly less pain. Ironically the researchers expected the steroid injection to be more efficacious. They had designed the trial on the assumption that the acupuncture needle was a 'placebo', and not expected to have an effect! By contrast Vickers and collaborators conducted a pragmatic controlled trial of acupuncture for headache in primary care where two groups were

---

⁷ Vickers et al. 2004.

randomised to a policy of either using or not using acupuncture.⁸ The acupuncture group received up to 12 sessions, with treatments based on traditional acupuncture principles, over a three-month period. At 12 months the acupuncture patients had significant and ongoing improvements in their headache symptoms. These two trials of acupuncture are illustrative: in the former the research design required a tightly controlled protocol, with treatment involving the use of a single acupuncture needle just once, an approach to treatment that did not reflect what patients would normally receive when consulting an acupuncturist; in the latter, the patients received routine acupuncture treatment and their experience within the trial was not modified by the study design.

Challenges ahead

Treatment flexibility

Within a pragmatic randomised controlled trial, there remain a series of challenges that need to be addressed. One of the challenges is deciding on what style of acupuncture should be assessed, and to what extent can or should this style be defined and described. The standardised approach to acupuncture that has been taught in the People's Republic of China since the 1950s, and often labelled 'Traditional Chinese Medicine' (TCM), has since been transmitted to the West after the opening up of China at the end of the Cultural Revolution. It has subsequently evolved into a number of expanded forms.⁹ While this is the most widely practised style in the United Kingdom, North America and Australia, there are other acupuncture traditions, some of which may be more appropriate for certain conditions. For example trigger point acupuncture has a particular focus on the treatment of musculo-skeletal pain, while the 'Five Element' schools of acupuncture, using uniquely modern interpretations of the canons of Chinese and Japanese acupuncture, have a reputation for treating mental and emotional problems.

---

⁸ Birch and Felt 1999.

⁹ MacPherson et al. 2004b.

Experimental trials will usually have highly constrained treatment protocols. They therefore are more appropriate in the case of acupuncture point prescriptions that are formulaic. In pragmatic trials there is scope for flexibility so that practitioners in the trial can deliver the sort of acupuncture that they normally would. In everyday practice, acupuncturists individualise their treatments for each patient, and for the same patient they ensure that treatment will follow the changes in the patient's condition over time.

There are two main options for defining treatment in a pragmatic trial. One approach involves practitioners recording carefully the key aspects of their treatments provided during the trial. Then after the trial this data is pooled so that as complete a picture as possible is presented. I have used this methodology in a large scale trial of acupuncture for low back pain.¹⁰ Treatment guidelines can then be developed after the trial. The second option is one where treatment guidelines are developed before the trial with the involvement of practitioners. One example of this method can be found in a trial of acupuncture for depression.¹¹ To develop guidelines for acupuncture treatment within clinical trials, a thorough literature review of good clinical practice is a useful starting point. Expert opinion can then be canvassed before drawing up a tentative set of guidelines. To add to their credibility, these guidelines could be subjected to a consensus process,¹² which would help firm up a comprehensive and acceptable framework. In this way practitioners should have sufficient flexibility to be able to address the expected variability in patients.

Patient variability

Given the limited resources available for acupuncture research, 'To what extent should the focus of research be narrowed to certain sorts of patients and to certain types of conditions?' While conventional medical research has usually focused on specific interventions for specific conditions, we know that patients who consult acupuncturists are often seeking help for multiple symptoms simultaneously.

Within traditional Chinese medical practice, different patients with

---

¹⁰ Allen, Schnyer and Hitt 1998.

¹¹ Black et al. 2001.

¹² Kaptchuk 1998.

the same disease can be diagnosed with different patterns of disharmony, while the same pattern of disharmony in different patients can be diagnosed for different diseases.¹³ Each pattern of disharmony or syndrome reflects a cluster of possible symptoms, and not all symptoms for each syndrome will be present in every case.

In addition there is wide variability in the patients' expectations, preferences, willingness to engage in a healing process, openness to change, scope for taking on supportive changes in their life-style, and so forth. All these factors can be expected to influence outcome. There is therefore a strong case for treatment guidelines to allow acupuncturists to engage patients as fully as possible in their healing process.

Broader outcomes

Another challenge is to establish the benefits of treatment. We need to consider carefully what sort of outcome measures to use, and when to apply them. Conventionally, trials often use a condition-specific measure. For example to measure depression there is the Beck Depression Inventory, in the form of a questionnaire with a series of questions related to depression. The responses are then converted into a validated scale which rates the severity of depression.¹⁴ In addition, trials usually measure the general health status or quality of life of the patients. An example of this is Short Form 36, which is a questionnaire covering eight subscales, including physical and mental functioning, bodily pain, and vitality.¹⁵ However from recent research we know that patients can experience changes with acupuncture that are not covered by these two types of measures.¹⁶ For example, patients can undergo changes in attitude, in behaviour, and in identity, any of which might be experienced as profoundly important. While short-term changes tend to impact on the disease-specific measures, these broader changes tend to have an impact over the longer term.¹⁷

---

¹³ Beck et al. 1961.

¹⁴ Brazier et al. 1992.

¹⁵ Paterson and Britten 2003.

¹⁶ Gould and MacPherson 2001.

¹⁷ Kaptchuk 1998.

Within the field of acupuncture research, therefore, there remains the challenge of identifying some way of capturing these broader changes, and assessing the residual gap between these and conventional outcome measures. There also remains the question as to the appropriate timing for taking these measures of outcome. Interestingly, some acupuncture trials are showing improvements in health well after the completion of treatment, so long term follow up seems a pre-requisite for future trials. With studies that are evaluating cost-effectiveness, the costs of acupuncture may only be offset by reduced conventional costs after a reasonable period of time, and hence this is another reason for choosing a longer term follow up.

Therapeutic relationship

A third challenge is for us to understand better the role of the therapeutic relationship. In much conventional 'evidence-based medicine' the therapeutic relationship is labelled a 'non-specific' component of treatment. This separation may make more sense in the context of a physician prescribing, say, antibiotic medication. The implication is that it does not have to be an acupuncturist that delivers the non-specific component. Some sceptics argue that the benefits of acupuncture are entirely due to the non-specific components, because acupuncturists working in European traditions of Asian medicine deliberately aim to develop good and lasting relationships with their patients, taking sufficient time and providing a friendly healing atmosphere.

However it can also be argued that the dynamics between acupuncturist and patient can be expected to be very specific to the context. For example, both the process of taking the case history and the way explanations are offered as to why a patient may be suffering from certain symptoms, are core components of treatment, without which certain outcomes might not be facilitated. Therefore a useful area of research is to clarify the way that much of what happens in the relationship between practitioner and patient is actually specific to acupuncture.

There is also a case to make that a good therapeutic relationship generates a synergy that can amplify the overall benefits.¹⁸ The inter-

---

¹⁸ MacPherson et al. 2003.

action itself has a healing component, such that in terms of outcome, one (the needling of certain points) plus one (therapeutic relationship) could equal more than two. While it is contentious whether separating out and then adding up the component parts of the therapeutic relationship is a worthwhile approach, support for the synergy argument comes from recent quantitative research that shows a direct association between patient-perceived empathy and enablement, as well as enablement and outcome i.e. the better the empathy, the more the enablement and the better the outcome.¹⁹ We need to be careful about our conclusions when reducing the full experience of the therapeutic relationship, which is so rich and varied, to its component parts. Indeed, establishing the potential role of synergy within a clinical trial would render less meaningful any evidence based on separating out the therapeutic relationship from the overall treatment process, as can occur in experimental controlled trials.

Qualitative methods

While clinical trials can tell us about average effects across a sample of patients, they tell us little about: individual experiences of treatment; the ways patients value the therapeutic relationship; the role of their input in helping themselves get better; their expectations and beliefs; and the outcomes that they see as particularly important. For a deeper understanding of these issues, that is beyond simple checklists of satisfaction and acceptability, we need to use qualitative methods.²⁰ These methods include structured observation, ethnography, action research, focus group discussion and in-depth interviews. While the appropriate qualitative method will be dictated by the research question, what I am arguing for is the need to develop a broader understanding of acupuncture and its impact that goes beyond numerical data, for example by providing detailed explanations of why people consult acupuncturists and what they value from their treatment. It is also possible to combine qualitative and quantitative methods, for example using in-depth interviews of patients

---

¹⁹ Ritchie and Lewis 2003.

²⁰ Paterson and Britten 2003.

as a nested study within a randomised controlled trial. In another example, Paterson explored changes to presenting symptoms in patients receiving a form of traditional acupuncture using both (quantitative) questionnaires and (qualitative) in-depth interviews. The qualitative data from interviews provided evidence that went beyond the quantitative data from questionnaires, with patients reporting whole-person effects characterised by changes in strength and energy, as well as personal and social identity.

Conclusion

When tackling research that aims to clinically evaluate acupuncture, there are a number of complex issues that need to be addressed. First, we require a reasonably robust methodology, one that can deliver results that are sufficiently rigorous to contribute to the developing evidence base. Second, pragmatic randomised controlled trials are more appropriate than experimental trials, in that they can answer real-world questions about the clinical and economic impact of acupuncture. Experimental controlled trials are more suited to theoretical questions concerning the impact of components of treatment, such as the needling *per se*; but they will tell us little about the provision of acupuncture in routine practice. Third, there remain a number of challenges in designing good quality and relevant trials. These include the design of a protocol that reflects routine practice with the flexibility required to match expected patient variability. Patients have reported broader outcomes from acupuncture, including changes to personal and social identity, and it is important to capture these changes, especially if they are valued as much or more than conventional outcomes measures. It will also be useful to identify those aspects of the therapeutic relationship that are seen to be of particular value to patients, and also establish whether there is synergy between the therapeutic relationship and other aspects of the acupuncture treatment. Given that people are turning to acupuncture in growing numbers, there is a need to respond to the increasing demands for good quality evidence, and to do this in a way that maintains the integrity of the acupuncture as a holistic treatment.

References

- Beck A., Ward C., Medelson M., Mock J., and J. Erbaugh 1961, 'An inventory for measuring depression', *Archives of General Psychiatry*, 4: 561–71.
- Birch S. and R. Felt 1999, *Understanding Acupuncture*, Edinburgh: Churchill Livingstone.
- Black N., Murphy M., and D. Lamping 2001, 'Consensus development methods, and their use in creating clinical guidelines', in *Methods of Evidence Based Healthcare*, Andrew Stevens et al. (ed.) London: Sage, 426–48.
- Brazier, J.E., Harper, R., Jones, N.M., O'Cathain, A., Thomas, K.J., Usherwood, T., and L. Westlake 1992, 'Validating the SF-36 health survey questionnaire: new outcome measure for primary care', *British Medical Journal*, 305.6846: 160–4.
- Garvey, T.A., Marks, M.R., and S.W. Wiesel 1989, 'A prospective, randomized, double-blind evaluation of trigger-point injection therapy for low-back pain', *Spine*, 14, 9: 962–4.
- Gould, A. and H. MacPherson 2001, 'Patient perspectives on outcomes after treatment with acupuncture', *Journal of Alternative and Complementary Medicine*, 7, 3: 261–8.
- Kaptchuk, T.J. 1998, 'Powerful placebo: the dark side of the randomised controlled trial', *Lancet*, 351.9117: 1722–5.
- MacPherson, H., Mercer, S.W., Scullion, T., and K.J. Thomas 2003, 'Empathy, enablement, and outcome: an exploratory study on acupuncture patients' perceptions', *Journal of Alternative and Complementary Medicine*, 9, 6: 869–876.
- MacPherson, H., Scullion, A., Thomas, K.J., and S. Walters 2004a, 'Patient reports of adverse events associated with acupuncture treatment: a prospective national survey', *Quality and Safety in Health Care*, 13, 5: 349–55.
- MacPherson, H., Thorpe, L., Thomas, K., and M. Campbell 2004b, 'Acupuncture for low back pain: traditional diagnosis and treatment of 148 patients in a clinical trial', *Complementary Therapies in Medicine* 12.1: 38–44.
- National Institutes of Health 1998, 'NIH Consensus Development Panel on Acupuncture', *Journal of Alternative and Complementary Medicine*, 280: 1518–24.
- Paterson, C. and N. Britten 2003, 'Chinese medicine acupuncture for people with chronic illness: combining qualitative and quantitative outcome assessment', *Journal of Alternative and Complementary Medicine*, 9: 671–81.
- Paterson, C. and N. Britten 2003, 'Acupuncture for people with chronic illness: combining qualitative and quantitative outcome assessment', *Journal of Alternative and Complementary Medicine*, 9, 5: 671–81.
- Ritchie J. and J. Lewis 2003, *Qualitative Research Practice: A Guide for Social Science Students and Researchers*, London: Sage.
- Roland, M. and D.J. Torgerson 1998, 'What are pragmatic trials?', *British Medical Journal*, 316: 7127: 285.
- Sackett, D.L., Rosenberg, W.M., Gray, J.A., Haynes, R.B., and W.S. Richardson 1996, 'Evidence based medicine: what it is and what it isn't', *British Medical Journal*, 312.7023: 71–2.
- Sauerland, S. and M. Maegele 2004, 'A CRASH landing in severe head injury', *Lancet*, 364.9442: 1291–2.
- Schwartz, D. and J. Lellouch 1967, 'Explanatory and pragmatic attitudes in therapeutic trials', *Journal of Chronic Disease*, 20, 8: 637–48.
- Vickers, A.J., Rees, R.W., Zollman, C.E., McCarney, R., Smith, C.M., Ellis, N., Fisher, P., and R. Van Haselen 2004, 'Acupuncture for chronic headache in primary care: large, pragmatic, randomised trial', *British Medical Journal*, 328, 7442: 744.
- World Health Organisation 2002, *Acupuncture: review and analysis of reports of controlled trial*. Geneva.