

# The clinical evaluation of traditional East Asian systems of medicine

**Hugh MacPherson, Karen Sherman, Richard Hammerschlag, Stephen Birch, Lixing Lao, Christopher Zaslowski**

## INTRODUCTION

This paper summarizes our exploration of the question. 'How best can we evaluate traditionally based East Asian systems of medicine?'. The use of randomized placebo-controlled clinical trials are considered the gold standard by the biomedical community and many funding agencies. Such studies are routinely, and appropriately, performed for the evaluation of new pharmaceutical drugs. However, the nature of traditional East Asian medicine, with its individualized diagnostic patterns and use of multiple modalities in a single treatment does not easily lend itself to this type of research design.

A first principle that has informed our deliberations is that the clinical evaluation should respect the integrity of the medicine. For example if we ask questions about treatment outcomes resulting from usual clinical practice, we would want the practitioners involved to be in reasonable agreement that the trial protocol adequately reflected what occurred in their normal clinical work. Developing an appropriate research question and using an appropriate design would be critical to this process. In this paper we will discuss a number of research designs and explore how well each works in the context of evaluating systems of East Asian medicine. While there is much current debate on how clinical evaluation of individual *modalities*, such as acupuncture, within traditional East Asian medicine can be both methodologically sound as well as sensitive to the underlying principles, the present paper extends this debate to research on *systems* of medicine.

## DEFINING A TRADITIONALLY BASED SYSTEM OF MEDICINE

In this paper, the term 'system of medicine' serves as a conceptual framework for a range of treatment modalities that share a similar or related theoretical orientation. For example both traditional Chinese medicine and Western biomedicine can be seen as systems in this context. Some commentators on the practice of Chinese medicine have criticized the tendency to label traditional Chinese medicine as a 'system'.<sup>1</sup> It has been argued that the word 'system' gives an erroneous impression of a rational, internally consistent and complete medical practice, thereby misrepresenting the diverse, complex and contradictory aspects of Chinese medicine. Therefore, in appreciation of the conceptual origins as well as contemporary practice of Chinese medicine, we broadly define a medical system as a set of evolving medical practices that may change over time and transform as a result of transmission across cultures.<sup>2</sup> If the system of medicine provides the broadest framework, then within the system a number of modalities may be incorporated. For example within traditional Chinese medicine the component modalities include acupuncture, moxibustion, Chinese herbs, exercise (*qigong*, *taijiquan*), nutrition and massage (*tuina*).

The modalities within a traditionally based system of medicine will share a similar or related theoretical orientation. They may also share some other characteristics, such as diagnostic procedures, treatment practices and explanations to patients. For

**Hugh MacPherson**,  
Foundation for Traditional  
Chinese Medicine, York, UK.

**Karen Sherman**, Center  
for Health Studies, Group  
Health Cooperative,  
Seattle, USA.

**Richard  
Hammerschlag**, Oregon  
College of Oriental  
Medicine, Portland, OR,  
USA.

**Stephen Birch**, Stichting  
for the Study of Traditional  
East Asian Medicine,  
Amsterdam, The  
Netherlands.

**Lixing Lao**,  
Complementary Medicine  
Program, University of  
Maryland School of  
Medicine, Baltimore,  
Maryland, USA.

**Christopher Zaslowski**,  
College of Traditional  
Chinese Medicine,  
University of Technology,  
Sydney, Australia.

Correspondence to:

**Dr Hugh MacPherson**,  
Research Director,  
Foundation for Traditional  
Chinese Medicine, 122A  
Acomb Road, York YO24  
4EY, UK. Tel: 01904 781630;  
Fax: 01904 782991; E-mail:  
hugh@ftcm.org.uk

traditional Chinese medicine, all modalities share underlying theories that include the concept of *qi* flowing through a system of interconnecting channels, disease as an imbalance in flow of *qi* or disharmony between *yin* and *yang*, and the restoration of balance through self-healing. Within the framework of each modality, there will be a range of techniques and procedures. For example within acupuncture as a modality, the selection of points for needle insertion might be guided by specific palpatory techniques or by *zang fu* diagnosis.

In evaluating a system of medicine, the component parts must be clearly defined. Moreover, in some circumstances, focusing the evaluation on a subset of modalities is appropriate. For example building on existing acupuncture and herb trials, these two modalities could be tested as a combination therapy. Needless to say, the system, or part of a system, under evaluation needs to be adequately defined with reference to literature and other sources.<sup>3</sup>

## ADDRESSING QUESTIONS OF CLINICAL EFFECTIVENESS

### Case and cohort studies

There are several kinds of studies that may be used in evaluating the effectiveness of traditional East Asian medicine. Since individualizing treatments at each visit of patients with the same biomedical condition is common, the case study has an obvious place among appropriate methodologies. For some researchers, as has been argued by Scheid<sup>1</sup> the case study has been the 'gold standard' in the evaluation of traditional Chinese medicine for two millennia. For mainstream researchers, the case study can be useful as background information on what a traditional system of medicine might successfully treat. When used in conjunction with a clinical trial, the case study can also provide depth to our understanding of the patient perspective, including their experiences of treatment and its wider influence on their health.

Another useful design is the cohort study, which would involve the systematic assembly of patients (either prospectively or retrospectively) using carefully defined inclusion and exclusion criteria like any other study. For prospective cohort studies focused on patient outcomes, one could explore how such factors as the condition, specific components of treatment, and specific characteristics of the patient and the practitioner, related to clearly defined outcomes. For example, when some patients do better than others, one may seek to draw conclusions

about which conditions, or which traditionally based diagnostic patterns, are associated with a faster recovery.<sup>4</sup> Such cohort studies can be a useful prerequisite to controlled clinical trials. Since funding is limited, controlled trials are best targeted at those conditions, diagnostic patterns or treatment algorithms showing most promise in cohort studies.

### Randomized controlled trials

In traditionally based systems of medicine, the interventions of normal clinical practice are extraordinarily complex and typically include using a diagnostic assessment that does not correspond with biomedicine's diseases and customizing treatment with many modalities. Therefore, randomized trial designs comparing packages of care are especially appropriate. In such studies, what is being assessed is the overall package of care,<sup>5</sup> and the treatment intervention is sometimes described as being within a so-called black box. This design allows the practitioners relative freedom within the black box to individualize treatment to each patient. This trial design, commonly known as the pragmatic randomized controlled trial, can provide a practical comparison where decisions need to be made between two therapeutic approaches, as when, for example, there are limited resources for the delivery of care.

To increase generalizability from pragmatic trials of traditional East Asian systems of medicine, a reasonable number of practitioners need to participate in the study. Consistency in assessment of the differential diagnosis and corresponding treatment between practitioners can be enhanced by special training prior to the trial. In addition, one may choose to establish minimum levels of consistency prior to the trial. It is usually possible within a pragmatic trial to build into the design some sub-studies which explore differences in outcome within the treatment group, such as those between practitioners, between patient sub-groups or between traditionally based diagnostic categories.<sup>6</sup>

When drawing conclusions about a system of medicine from the results of a pragmatic trial, a detailed description of the interventions is necessary to facilitate understanding the results of the trial. If non-biomedical diagnostic assessments are used to determine treatment, then a breakdown of the diagnostic assessments and their relationship to treatments performed is necessary. The range of points, the auxiliary treatments such as moxa and cupping, the prescribed self-help activities and lifestyle advice must all be described in a manner that facilitates reproducibility. One way to do this is by creating a pragmatic trial manual, so that others can better understand the trial and also reproduce the trial to

extend generalizability and broaden the evidence base.

While the pragmatic trial is an obvious choice for evaluating a complex intervention, the more structured approach of the explanatory trial can examine the specific effects of the therapy by controlling for all the non-specific effects. In the context of a complex intervention, where treatments are typically customized for patients according to East Asian medical diagnoses, the use of a treatment manual to guide diagnostic and treatment decisions, and thus permit reproducibility, may be important in such explanatory trials.<sup>7</sup> To develop such a manual, which should be done before the trial is undertaken, of the use of a comprehensive approach that includes literature reviews, practitioner treatment data, expert panels, and consensus among practitioners would be important. The specified protocol must then be followed by all practitioners in the trial itself. With an explanatory trial, the choice of control group depends on a number of factors, including the research question and the specific condition. On this basis, the control group would probably receive one of many possible invasive or non-invasive sham acupuncture procedures.

While the focus of this paper has been on systems of medicine, it may also be of interest to use an explanatory trial to evaluate the combination of two or more modalities within a system. For example one could test acupuncture plus herbs vs sham acupuncture plus placebo herbs. Another option for teasing out modalities within a system would be a trial comparing two modalities from the same system, for example acupuncture versus Chinese herbs. Here one could use a 'double-dummy' approach to masking patients to treatment intervention, where one group of patients receives real acupuncture and placebo herbs, while the other group receives non-invasive or invasive sham acupuncture and real herbs.

## USING TRADITIONALLY BASED DIAGNOSTIC CATEGORIES

One of the challenges of a traditionally based medical system is that patients with one condition, say low back pain, may have different underlying patterns of disharmony, each of which is treated differently in normal practice. Uniform treatment for all patients would be expected to lead to sub-optimal results. Thus, targeted treatment for each pattern would be desirable, if such patterns can be reliably diagnosed. For example, a trial of patients with low back pain may find that patients' patterns of disharmony could be grouped in three traditional Chinese

medical categories: *qi and blood stasis, cold and damp channel obstruction syndrome and kidney (yin or yang) vacuity*.<sup>8</sup> Customizing the treatment to these patterns, rather than providing standardized treatments, would both reflect normal practice as well as be more likely to deliver better results.

However if treatment is customized towards such categories, it is important to ask how reliable practitioners are in identifying the patterns of disharmony. One approach to this question of reliability is to provide pre-trial training to all participating practitioners, ensuring that a minimum level of reliability is achieved as a pre-requisite. Such an approach was used in a trial for the treatment of depression with acupuncture.<sup>7</sup>

Another approach is to focus only on one pattern of disharmony, one that providers believe they can treat successfully with traditional East Asian medicine (i.e. enabling a smaller sample size) or one that is particularly common (i.e. facilitating recruitment). To increase the potential effect size in a trial, one would 'double-screen' patients prior to randomization.<sup>9</sup> The first screen would be for a biomedically well-defined condition, such as migraine, and the second screen would be for the traditionally based category, e.g. *ascendant hyperactivity of liver yang*. The use of a double screen could simplify the protocol and the interpretation of the results. In this situation, it would be critical that practitioners agree on how to categorize the patients. After both assessments, patients would be randomized to two groups, one of which would receive treatment targeted for that specific diagnostic category, while the other group would receive a control treatment. The actual care the control group received would depend on the research question and other factors such as the patients' condition.

In addition, we recommend evaluating the potential of outcome measures specifically designed for traditionally based diagnostic categories. For example, as well as measuring changes of frequency and intensity of pain associated with migraine headaches, it may be possible to design and validate an outcome that includes assessment of change to the wider range of symptoms associated with the underlying pattern. For *ascendant hyperactivity of liver yang*, for example, such an outcome measure could assess not only pain, but also emotional state, such as irritability, physical sensations, such as heat, and relevant signs, such as from the pulse and tongue.

## CONCLUSIONS

The clinical evaluation of traditionally based East Asian systems of medicine raises some important

issues for researchers and practitioners. A central concern is that the methodology adopted for such evaluations does not adversely impact the quality of treatment. To support the integrity of the 'system', which may incorporate a complex array of interventions, choosing the right research question and the related methodology is a crucial starting point. In this paper, we have outlined a number of research designs that are well suited to traditional East Asian medicine. Individual case studies and cohort studies have an important role in understanding and improving practice. For definitive answers on effectiveness, the pragmatic randomized controlled trial can be both methodologically rigorous as well as providing results relevant to normal clinical practice. For establishing an effect of treatment per se, over and above the non-specific effects, then the more constrained explanatory controlled trial also has a useful role. However our underlying concern in this paper is that sensitivity to the principles and practices within the East Asian systems of medicine must inform our approach to clinical evaluation.

#### REFERENCES

1. Scheid V. Orientalism revisited: reflections on scholarship, research and professionalisation. *Eur J Oriental Med* 1993;2:22–33.
2. Scheid V. Chinese medicine in contemporary China: plurality and synthesis. Duke Univ Press, Durham 2002, in press.
3. Birch S. Testing the claims of traditionally based acupuncture. *Complement Ther Med* 1997;5:147–151.
4. MacPherson H., Fitter M. Factors that influence outcome: an evaluation of change with acupuncture. *Acupunct Med* 1998;16(1):33–39.
5. MacPherson H. Out of the laboratory and into the clinic: acupuncture research in the real world. *Clinical Acupuncture and Oriental Medicine* 2000;1(2):97–100.
6. Thomas K.J., Fitter M., Brazier J., MacPherson H., Campbell M., Nicholl J.P., Roman M. Longer term clinical and economic benefits of offering acupuncture to patients with chronic low back pain assessed as suitable for primary care management. *Complement Ther Med* 1999;7(2):91–100.
7. Allen J.J.B., Schnyer R.N., Hitt S.K. The efficacy of acupuncture in the treatment of depression in women. *Psychol Sci* 1998;9:397–401.
8. Birch S., Sherman K. Zhong yi acupuncture and low back pain: traditional Chinese medical differential diagnoses and treatments for chronic lumbar pain. *J Altern Complement Med* 1999;5(5):415–25.
9. Lao L., Ezzo J., Berman B.M., Hammerschlag R. Assessing clinical efficacy of acupuncture: considerations for designing future acupuncture trials. In: Stux G., Hammerschlag R. eds., *Clinical acupuncture: Scientific basis*. Springer-Verlag Berlin. pp 187–210, 2001.

#### CONTRIBUTORS

An international group of experts in acupuncture research, who met July 2–4, 2001 at Exeter University, Exeter, UK, helped develop the ideas that form the basis of this paper. In addition to the co-authors, participating in this group were: Mark Bovey, Acupuncture Research Resources Centre, University of Exeter, UK; Sarah Budd, Acupuncture Research Resources Centre, University of Exeter, UK; Val Hopwood, Acupuncture Association of Chartered Physiotherapists, Coventry, UK; George Lewith, University of Southampton, UK; Kenji Kawakita, Meiji University of Oriental Medicine; Marco Romoli, Federazione Italiana Società di Agopuntura, Prato, Italy; Kien Trinh, McMaster University, Canada; Adrian White, University of Exeter, UK.