Measuring the efficacy of acupuncture medicine

Dr. Johannes G. Schmidt
Family Practice & Institute for Clinical Epidemiology,
Foundation «Paracelsus Today», Einsiedeln/Switzerland

Randomised controlled trials about the efficacy of Traditional Chinese Medicine (TCM) conducted so far show a number of methodological shortcomings. A vital problem is the definition of homogeneous study groups, using nosological criteria of conventional medicine together with nosological criteria of TCM. The placebo effect is a secondary question which should not distract from the primary question of the efficacy of a complete therapy and its setting. The potential of TCM *in not only treating the actual symptoms but in preventing a multi-morbid chronic patient career should become a major research questions of future studies.*

Traditional Chinese medicine (TCM) such as *acupuncture* is often rejected by conventional «Western medicine» purely because of ideological prejudice. However, it cannot be overlooked that there is also a lack of good quality research on the efficacy of *acupuncture medicine* which makes it difficult to document the value of this therapeutic school.

Clearly, *case histories* and *case series* are inadequate proof of therapeutic efficacy, although such observational evidence may help to define valuable hypotheses. The need for properly conducted controlled trials is often insufficiently understood, not only in unconventional medicine, but in conventional medicine as well. Let me illustrate this with an example: doctors dealing with the treatment of patients after myocardial infarction may observe that the presence of arrhythmia is associated with an impaired prognosis, patients with arrhythmia dying obviously faster than patients without arrhythmia. Moreover, patients with arrhythmia successfully suppressed by the use of drugs appear to have a much better prognosis than less successfully treated patients. *Case series* would thus suggest a clear benefit of the anti-arrhythmic treatment. The true story, however, is that antiarrhythmic treatment in patients after myocardial infarction is clearly harmful as it increases mortality dramatically. This was eventually shown after properly controlled trials had been conducted.

This example shows two important things. **First**, observations from case histories can lead to gross errors. **Second**, the research paradigm of conventional medicine is changing from *mechanistic speculation* to *controlled empiric observation*, exemplified by the development and growing importance of *clinical epidemiology*.

The major feature of this development is the *critical questioning of what really matters for the patient* and *whether treatment and research findings are relevant* and applicable to practical medicine. It is becoming increasingly accepted that *the studied treatment must reflect the treatment situation in practice* and must not be distorted by the need of standardisation for study purposes.

Modern research methodology - and I must emphasise that much of the research within conventional medicine is also still invalid in these terms (as other contributions in this book well illustrate) - modern methodology is not incompatible with research needs in unconventional medicine, and I believe that these methods should and can be used for the study of *TCM therapy*.

A number of *randomised controlled trials* have attempted to assess the efficacy of *acupuncture* therapy. These studies, however, show a number of methodological shortcomings. Most of the studies were too small to have adequate statistical power. Many studies achieved inadequate follow-up, preventing a methodologically valid intention-to-treat analysis. Often, also, the period of follow-up was too short to assess important endpoints. A vital problem, hardly resolved in studies conducted so far, is the definition of homogeneous study groups, using nosological criteria of conventional medicine together with nosological criteria of *TCM*. The many studies using a *fixed formula acupuncture* can hardly be regarded as methodologically valid because they study a distorted treatment which does not exist in normal practice.

In planning a trial, what is an active treatment and what is a control?

It is quite impossible to conduct *double-blind* trials with *acupuncture* because an acupuncturist can never treat blindly. The double-blind design may perhaps be applied if a pure herbal treatment without acupuncture or moxibustion is indicated and if a truely inert placebo tea of similar taste can be provided. A placebo controlled study of Chinese herbal therapy in atopic dermatitis, for example, has shown some improvement during a standard-formula herbal therapy, but as the authors acknowledge, the study is of limited value because important principles of *TCM*, i.e. using an individualised formula, were not adopted.

Single-blind studies may perhaps provide some information about the magnitude of the *placebo* effect of *TCM*, but they suffer from other shortcomings. *Sham acupuncture*, for example, may not be inert and seems to be an unnecessarily complicated procedure. No one, in comparison, would consider sham surgery as a control for the study of the efficacy of a surgical procedure although surgery, no doubt, may have *placebo* effects.

The *placebo* effect, perhaps, may be produced by nothing more than an unspecific influence on the *qi* flow, and the therapeutic setting of *TCM*, therefore, may enhance this practically valuable effect. The distinction of a *placebo* effect can be of some (perhaps mainly academic) interest, but for a patient seeking cure, it doesn't matter through which known and unknown mechanisms he or she is helped.

Consequently, if, for technical reasons, a *placebo* effect cannot be properly controlled without distorting the practice of *TCM*, we should not let ourselves be confused. The *placebo* effect is a secondary question which should not distract from the primary question of the efficacy of a complete therapy and its setting. In fact, this logic was always applied if a true *placebo* was not feasible for practical reasons, exemplified by the *lack of placebo-surgery* and by a large number of studies using «*referred care*» or «*usual care*» as controls.

Practically relevant research on the efficacy of *TCM* may best be designed as a comparison of *TCM* with conventional medicine by imitating the pragmatic procedure of comparing a new treatment with a standard treatment. Often, accepted standard treatments may in fact not be superior to a *placebo*. For such pragmatic reasons we should accept *TCM treatment* only as «successful» if it is significantly **better** than conventional standard treatments. Therefore, the study hypothesis must be that *TCM is better than conventional treatment* (or, as some may like it, the *null-hypothesis of no difference* should be refuted). Thus, one should accept that *TCM* is worth being introduced into medical practice only if it is shown to be superior for a given condition (or perhaps if it is more cost-effective). By accepting this prerequisite, the use of conventional treatment as a control is certainly a valid method (as already mentioned, such a control is often termed *referred care* or *usual care*).

How should study subjects be selected?

A proper study must have an *internal validity* and a meaningful *external validity*. The overriding issue of *external validity* is that the patient group from which the data stem are generalisable. The conditions which are studied should be narrowly defined, because otherwise it will be difficult to delineate the exact disease condition to which the results apply. For example, if a study about the treatment of headache showed *TCM* to be successful, it would of course be important to know for which type of headache this is the case.

There is, however, a *conventional Western* and a *Chinese nosology* of headache. Conventional medicine classifies headache into *migraine*, *tension headache*, *cluster headache*, *posttraumatic headache*, etc.; for *Chinese medicine* headache can occur

due to *rising liver yang*, *liver* or *stomach heat*, *wind*, *dampness* or *phlegm*, *deficiency* or *stagnation of qi*, etc. Moreover, there are acute and chronic forms and a great variability in frequency and duration exists. It is therefore quite questionable whether a study population fulfilling the criteria of a Western diagnosis (say *tension headache*) is homogeneous at all. Defining a strictly homogenous study population may therefore be futile.

However, if *entry criteria* and *exclusion criteria* are of practical relevance in the study and clearly delineated at the same time, the study population is sufficiently delineated, even though some heterogeneity of the study population may persist. We must bear in mind that this does not necessarily impair the *internal validity* of a trial, if the study and control group are properly randomised. If the *size of the study* is adequate, the comparability of study and control group is guaranteed by a proper *randomisation procedure*, where *random errors* are accounted for in the eventual *statistical test of significance*. However, it is of importance to assure a thorough evaluation of patient *baseline characteristics*, including *Chinese* diagnostic criteria. Only this will allow the comparison of the study with other studies (*external validity*), and to some limited extent, the analysis of subgroups or *confounders* and *prognostic variables*.

The *internal validity* is also dependent on effective *limitation of bias*. Therefore, *complete follow-up* and an *intention-to-treat analysis*, *blindness* in the assessment of the *treatment outcome*, the avoidance of a treatment contamination between both arms, a correct and dedicated management of the treatment in both arms of the study, which may need special training and supervision, all these are essential ingredients of a valid study. Trials should be planned and conducted together by advocates and sceptics of *TCM*.

What is treatment success?

This is perhaps the most vital question of modern research. What are meaningful clinical variables? The selection of study *endpoints* must not be reduced to effects which are easily measurable but perhaps little relevant. In fact, the *hard data dogma* and its spurious validity have increasingly been questioned by researchers in conventional medicine in a large body of literature (see contribution by ALVAN FEINSTEIN, p. 210). Therefore, it may be wrong if practitioners of unconventional medicine, attempting to do research, believe that good research is mainly a problem of defining *hard, «objectively measurable» data*.

Furthermore, a range of complex measures that take the overall *quality of life* into account have been proposed. Therefore, it may also be wrong to believe that the wheel of patient-oriented outcome measures still remains to be invented for

research in unconventional medicine. In assessing treatment effects on headache, for example, *measurement scales* incorporating the subjective response by the patient, the amount of pain-killers used and the extent of disability in daily activities can be applied.

However, as many practitioners of *Chinese medicine* claim a long-term benefit of their treatment in preventing not only detribution from the original disease, but also in preventing the development of affliction of additional organ systems, the *entire multi-morbidity* may be a practically very important long-term endpoint. Headache for example, depending on the *traditional Chinese nosology*, may be associated with or, if untreated, develop *chronic fatigue*, *anemia*, *sleeping disorders*, *hypertension*, *stroke*, *chronic tonsillitis* or *sinusitis*, *spastic disorders* or *tremor*, as well as other multi-morbid complaints and conditions. In *TCM* terms, the condition leading to headache is considered untreated if merely drugs to relieve the pain are used as in conventional medicine.

I believe that the potential effect of *TCM* in not only treating the actual symptoms and disease but also in the prevention of a multi-morbid patient career should become major research questions of future studies about the efficacy of *TCM*. *Acupuncture and TCM* are only more useful and attractive than conventional pain-killers and other treatments if it can be shown by good quality research that it makes a difference in the long-term perspective. This requires long and relatively large studies, but there is no other way if we want to know the important questions about *TCM*.

Further reading

LEWITH GT, ALDRIDGE D (eds). Clinical research methodology for compelementary therapies. Hodder and Staughton, London - Sydney - Auckland 1993

TER RIET G, KLEIJNEN J, KNIPSCHILD P. Acupuncture and chronic pain: a criteria-based meta-analysis. *Journal of Clinical Epidemiology* 1990; 43: 1191 - 1199

PATEL MS. Problems in the evaluation of alternative medicine. *Social Science and medicine* 1987; 25: 669 - 678