Research methodological issues in evaluating herbal interventions

Abstract: Randomized controlled trials provide the best evidence, and is seen as the gold standard for allopathic research. Herbal therapies are not an integral part of conventional care although they are still used by patients in their health care management. These medicines need to be subjected to rigorous research to establish their effectiveness and safety. Clearly defined treatments are required and should be recorded in a manner that enables other suitably trained researchers to reproduce them reliably. Quality control of herbal products is also a prerequisite of credible clinical trials. Methodological strategies for investigating the herbal interventions and the issues regarding appropriate patient selection, randomization and blinding, placebo effects and choice of comparator, occupational standardization and the selection of appropriate study endpoints to prove efficacy are being discussed. This paper will review research options and propose some suggestions for future research design.

Keywords: CAM research, herbal therapies, methodology, clinical trial

Introduction

Herbal medicines, or phytomedicines (to give them their modern European name), are closer to conventional drugs than other complementary and alternative medicine (CAM) approaches. In many countries herbal medicines are even prescribed by doctors alongside modern drugs and dispensed or supplied primarily by pharmacists. It is only in the USA among the developed countries that legislation has established herbs as ‘dietary supplements’ and thus removed them formally from the medical scene. Herbal medicines consist of many chemical constituents with complex pharmacological effects on the body. They have been used continuously for many decades even centuries, often in ways much different from those of modern medical prescribing. The widespread use of herbal medicines suggests, though does not assure, the safety and efficacy of these medicines. The lack of pharmacological and clinical data on the majority of herbal medicinal products is a major impediment to the integration of herbal medicines into conventional medical practices. There is often serious resistance in meeting the research requirements of legislators. Evaluation of efficacy of herbal products and applying the principles of modern medicine is a paramount issue. The “gold standard” of evidence for treatment efficacy within evidence based medicine (EBM) is considered to be a systematic review or meta-analysis of randomized controlled trials (RCTs) with double blinding and a comparator group. The evaluation of herbal products with similar principles of EBM poses certain unique issues. While new clinical trials are being published at an increasing rate, there are still huge gaps in the evidence base and
most clinical studies are judged as not suitably rigorous.\textsuperscript{5} The present review discusses the issues regarding the designing of clinical trials of herbal products using the principles of modern medicine. These methodological challenges are not unique to herbalism and similar principles can be applied to other forms of CAM.

**Why research into effectiveness of herbal medicines is important?**

Herbal products have always been an important part of the public’s healthcare around the world.\textsuperscript{2} Surveys of CAM therapies have shown the wide usage of these products.\textsuperscript{3} The herbal products may have many ingredients, often with varying concentrations of the therapeutic compounds between products and between different batches of the same product. The issue of quality control and the selection of appropriate dosage regimens have been emphasized. A single formulation and dosage form with maintained consistency in multiple batches should be used throughout the different stages of the clinical trials.\textsuperscript{4} Although traditional complementary medicine and its practitioners have not demanded clinical trials such clinical trials are a requirement for modern scientists. As the use of herbal products rises, clinical investigation of these practices becomes increasingly important. This is because once the efficacy is proven, alternative treatments can be endorsed.\textsuperscript{5} Difficulties abound in determining efficacy due to the variety of methods used to compare different therapies and lack of comprehension of the model of holistic medicine.\textsuperscript{5}

**What are the challenges?**

In the past herbal therapy research has been shown to have many methodological shortcomings which raise serious concerns regarding the efficacy and safety of the interventions studied.\textsuperscript{6} There is a hint of double standards in the way the clinical trials of herbal products are considered in comparison to modern medical investigation.\textsuperscript{7} However, there are numerous examples in the medical literature of herbs studied in the context of Western medicine (eg, several clinical trials of St. John’s Wort, used in the treatment of depression have been published using the principles of modern medicine).\textsuperscript{8} Herbal research does require special attention to “model validity”\textsuperscript{9} or the likelihood that the research has adequately addressed the unique diagnostic taxonomy and therapeutic context of the CAM system under investigation.\textsuperscript{9}

The issues that really need to be explored include issues that are related to the financial, ethical, product standardization (quality control), design of study and the regulatory requirements before filing an investigational new drug (IND) for conducting large phase III trials. WHO (World Health Organization) recently issued an operational guideline regarding these regulatory requirements needed to support clinical trials of herbal products.\textsuperscript{10} The evaluation of research quality in herbal medicine uses the same approach as that in the clinical trials of modern medicine, with additional components relevant for herbal products. Numerous “quality rating” systems have been developed for the evaluation of clinical research most notably the CONSORT (Consolidated Standards of Reporting Trials) group has produced a widely adopted set of reporting guidelines for RCTs of herbal medicine.\textsuperscript{10,11}

Data from one clinical trial is generally not sufficient to provide adequate information regarding treatment efficacy, safety, or cost benefit. Rather, a focused evaluation of a specific topic must be done, with data pooled from several clinical trials and then analyzed and interpreted together. The relevance of meta-analysis and systematic reviews of CAM are that they can become an acceptable format for including and determining which particular research studies accurately focus on specific parameters.\textsuperscript{12} Meta-analysis and systematic review can lead to better generalizations about treatment outcome.\textsuperscript{12} Recently, the systematic review of herbal medicines in a specific disease and of specific herbal medicines have been published.\textsuperscript{13–15} The risk-benefit profile of a few herbal medicines has also been reviewed based on the available evidence.\textsuperscript{16}

**Patient selection**

For a study to qualify as clinical research, a clearly defined treatment and indication are required. The inclusion criteria can be based on modern medicine or herbal medicine diagnosis. As the two systems are based on different approaches to understanding health and disease, the disease criteria may be different in some cases.\textsuperscript{17} However, it becomes increasingly difficult to be sure of generating a homogenous group of participants with a diagnosis defined by herbal medicines. For example, a homogenous group of osteoarthritis patients may represent more than a dozen diagnostic types when evaluated by traditional Chinese medicine (TCM) and each requiring a different treatment.\textsuperscript{17} A standardized approach, in which a treatment approximates the average syndrome, is often the way investigators handle this situation. Such an approach has advantages in that it simplifies the treatment strategy and often allows for the maintenance of blinding and prevents treatment confusion between groups. The main disadvantage is that it may result in suboptimal patient treatment and so
produce false negative results. Bensoussan and colleagues conducted a study with poor model validity that illustrated this problem. In a series of patients with irritable bowel syndrome treated with traditional Chinese medicine, one third of the subjects were treated with a standardized approach, one third with a treatment that fit the TCM model of medicine, and one third served as controls, receiving placebo. Subjects treated with a treatment that fit the TCM model improved for longer than those given the standardized approach, although the latter was easier for research purposes. Both groups did better than the placebo control group.

Weigent and colleagues proposed that selection can be based on modern medicine or the complementary medicine system alone. Another way to approach different diagnostic taxonomies of alternative medical systems is to provide for double classification and selection in a clinical trial. A double classification approach is possible in many situations, although it may add considerable complexity to the study. The “double classification” method is where subjects primarily meet Western diagnostic criteria and then are further classified according to the CAM system. Treatments are delivered according to CAM classification and outcomes are evaluated by criteria for both systems. If a traditional indication is being tested it is advised to mention the underlying traditional theories and concepts in the protocol. The key requirement is to have stringent inclusion and exclusion criteria so that the indication as well as treatment can be assessed reliably.

Placebo and nonspecific effects
CAM practices provide psychosocial support integrated with physical treatments and so make full use of placebo effects in the sense of maximizing nonspecific factors essential to any successful practice. The patients play an active part in the outcome of all the treatments in herbal medicine. The patients’ motivation to follow the treatment regimen is likely to be influenced by their preferences before treatment is started. The greater the need for participation of patients, the greater the motivation will influence the outcome. Individual expectation and the nature of the information given to patients (including the informed consent procedure) can impact outcomes significantly. In the herbal medicine, most treatments are complex, consisting of a mix of active components with contextual factors such as the way the treatment is taken. These contextual factors influence the patients’ perception and motivation towards the therapy and can also influence the outcome. The interaction of these specific and contextual or placebo effects is an important area for focused research.

It has also been argued that the artificiality of the RCTs may reduce the placebo element of any intervention, failing to capitalize on the nonspecific treatment effects, and therefore the trial will inevitably reflect the minimum level of benefit that can be expected. However, it is possible to apply RCTs without adopting a reductionist approach. These contextual factors can be equalized in the two groups by appropriate randomization and maintaining blinding.

Randomization
The comparability of two or more groups can be ensured by randomization so that any extraneous variables have a similar effect on all/both groups. Many herbal experts emphasize joint decision-making and pay due consideration to patient preferences. Since the practitioner is an integral part of the therapy, this process can be judged to be interfering in the treatment process. There is also a belief that randomization can interfere with a therapy (by eliminating choice) and obscure our awareness of how bias affects it. For example, traditional Chinese science is based not on linear cause-and-effect assumptions but on an assumption that “correspondences” occur between system levels (biological, social, cosmic). Despite these concerns, numerous herbs have been tested in randomized controlled trials. Some ways have also been suggested. Brewin and Bradly who proposed an approach called partial randomization whereby patients are given the choice to indicate a treatment preference. Those who express a preference are given their chosen treatment and the rest are randomized. Zelen also put a somewhat similar solution to this difficulty. Each patient was randomized to one of two groups: a do not seek a chosen treatment consent group who receive standard treatment or a seek a chosen treatment consent group who are asked if they would be willing to receive the experimental treatment. Some of the patients may decline the new treatment in which case they will receive standard treatment. The assumption is that the consent of patients need only be sought if they are randomized to the seek consent groups. The other patients will receive the treatment they would have received if they were not part of the trial.

The major limitation with these designs is that the patients who agree to be randomized may not be typical and ethical concerns of not asking consent are also there. A double screening recruitment model has been suggested to allow the practitioner to individualize treatments and to ensure the reliable assessment of therapy. This is similar to double classification proposed earlier. The subjects are first screened according to modern medicine diagnosis and then
categorized according to traditional classification system. The categories proposed can be defined into strata and patients can be equally randomized to the treatment and placebo arm.34

Design adaptive allocations have been suggested as an alternative to randomization that virtually guarantee a better balance than randomization, both for measured and unmeasured outcomes. This is especially important when resources are scarce, such as in herbal medicine trials and where underpowered studies might lead to premature termination of promising research paths.35

**Blinding**

Blinding is an important element of RCT as it helps to isolate the placebo effect and the observer bias. RCTs in medicine are usually conducted double blind which means neither the prescriber nor the patient knows about the treatment allotment. This ensures that data collected are unbiased and the attitudes of patients do not influence their response. In many cases of herbal medicine trials it is often impossible to maintain double blind. This becomes the case when the treatment is multifaceted, involving counseling, listening, explaining, lifestyle and dietary advice as well as prescribing herbal medicines. However, such interventions can be single blinding, where the therapist but not the patient knows the treatment allocation.

Another proposed approach to ensure blinding is that the therapist does not assess the outcome of treatment. The assessment is done by a third person who, like the patient is unaware of the actual intervention received. The cautionary principle is to ensure that the patient does not give the assessor details of the treatment that might unblind the assessor. The inclusion of self-report inventories to assess the outcome is another suggested approach.36

**Choice of control or comparator intervention**

RCTs often depend on matching the control group as closely as possible with the intervention group. If the trial is intended to provide the evidence of a specific effect of herbal medicine it is important that the comparator is similar to the treatment. Factors that should be standardized include color, odor, duration, frequency of intake, and credibility of the treatment to the patient and physical situation in which treatments are administered.37 Evaluation of complex natural products that have a distinctive odor, such as ginger, have posed challenges in selection of matching controls.38 Here, the concern was that a control intervention having no odor might elicit suspicion while a distinctive odor might exaggerate any potential effect.

There have been some trials reporting the use of placebo that are matched to color and smell of the active intervention, but contains no active ingredients.38

Trials of an herbal decoction requiring preparation at home also pose a great challenge in selecting the control. If the control group contains active ingredients then the comparison between experimental group and control group will be affected. A few convincing trials have been performed in European countries on the herbal treatment of atopic eczema.39–41 Clinical trials were performed in children with moderate to severe eczema and adults with refractory widespread atopic dermatitis. The herbal decoction prepared contained ten herbs. Fairly convincing evidence of efficacy was presented from these placebo controlled trials. However, problems involved the unpalatable nature of the decoction and difficulties encountered with preparing the decoction at home. Though it may be challenging to construct a matching placebo for certain herbal products, it is not impossible.

During the conduct of active comparator trials, the major consideration is the selection of the comparator drug. Whether it should be a herbal product or a modern medicine needs consideration.

**Sample size**

In conducting an appropriately powered study, sample size needs to be calculated. To analyze the sample size the stated effect size needs justification. In the a study by Bian and co-workers which evaluated 167 studies the authors found that only 0.6% RCTs conducted a priori sample size calculation.42 An increasing number of journals are requiring RCTs to describe this process in accordance with the CONSORT statement.43 The estimation of effect size and variance between groups from reported literature or expert opinion should reflect both clinical acumen and be of potential public health effect. The effect size and variance can vary widely between studies already reported in literature and the investigators are advised to conduct pilot trials with the test formulations to adequately estimate sample size. The clinical investigators must participate in this process of sample size calculation along with the statistician.44

**Minimizing therapist and protocol variability (Occupational Standardization)**

In the literature on psychotherapy, a number of studies have shown that the experience of the therapist is related to the outcome of the treatment.35 It is therefore essential that therapist have appropriate training in the herbal medicine
under investigation. The need to maintain occupational standardization has been realized.\(^6\) This is more evident in trials of herbal products where the therapist has an active role to play.

Additionally it is crucial that the treatment is delivered according to the approved protocol. Holistic practitioners may feel uncomfortable delivering standardized care. They may like to provide best therapeutic practice and thus protocol violations may occur. This therapist variability can be reduced by providing treatment manuals detailing the precise procedures to be followed.\(^7\) Another approach to control for therapist variability is to employ more than one therapist.\(^8\) Each therapist should treat similar number of patients. If the number permits, the results of the two therapists can be compared.

**Evaluation of outcomes**

The treatment must be recorded in a manner that enables other trained researchers to reproduce it reliably. This often requires objective endpoints. Herbal experts utilize a system of clinical observations which today might be considered obsolete and over subjective. Modern clinical trials insist on having data with hard endpoints that can be monitored. Thus there is a need to develop the means to objectively assess the subjective signs.\(^17\)

Herbal experts may not accept the traditional outcome measures such as symptom relief, cure or survival. Rather, they wish to evaluate outcomes in terms of alterations in energy balance, healing, ability to cope with problems. The effect of treatment on quality of life (QoL) has been emphasized and more modern medicine clinical trials are including questionnaire assessing the effect on QoL. Clinical researchers are challenged to develop more comprehensive, clinically relevant, and patient-centered research designs and outcome measures that will assess the diverse needs and values of populations.\(^7\)

**Assessing the effects of individual differences**

The first argument against the applicability of RCTs into complementary therapies including herbal medicine is that these are often very individualistic in approach and cannot always be standardized as a treatment for a large group of individuals in the context of an RCT.\(^7\) Moreover, the expectations and strong beliefs toward herbal medicine can influence the outcomes in trials. Baseline assessments of various psychological factors such as personality and mood must be carried out. These can be used as prognostic variables and can be used to check that randomization has produced comparable groups.\(^21\) The between treatment differences in outcomes due to expectations can be assessed by giving detailed information about the treatments prior to randomization and their expectations of effectiveness of each treatment assessed for example by means of VAS (visual analogue scale).

Certain other research issues that have been observed regarding clinical trials of CAM (that may not strictly be applicable to herbal products only) are the justification about the onset and duration of effects in the absence of preliminary studies. It has been realized that study designs often fail to specify and justify an optimal time for conducting follow up.\(^21\)

Another concern is the problem of drop-ins and sample contamination. This problem arises because many herbal products are available outside the medical setting as dietary supplements. The investigators should monitor their trial subjects whether they seek out herbal products outside the specified protocol.\(^21\)

**Conclusion**

In conclusion, outcome measures should reflect the intervention and indications tested. The primary end point should be unique, focused and specific and can be a “hard end point”. The secondary end points are supplementary and may include changes in QoL, mood, patient preferences. These endpoints become more important when the primary endpoints are expected to change slowly and would be particularly important when chronic problems are faced.\(^7\) Thus when performing RCTs, testing herbal interventions, concepts that go beyond western medical technology and understanding may be relevant. There is a need to develop and validate a range of measures applicable to modern medicine and herbal medicine.

**The future**

Whether conventional healthcare is prepared to accept complementary medicine or vice versa is debatable as one challenges the autonomy of the other. The specific effects of the therapies, how they should be used and delivered to optimum benefit, need to be established. Integrated Medicine. The Way Forward for the Next 5 Years (Foundation for Integrated Medicine 1997) proposed an examination of the research issues of efficacy, safety, biological plausibility, methodology, and funding.\(^49\)

In brief, good research means better patient care. Instead of simply condemning the RCT out of hand, it has been seen that herbal products can be evaluated using an RCT approach, albeit in a modified form.
There is still a hunger for beneficial natural medicines and this has renewed interest in the potential of indigenous medical traditions. Standards can be developed for herbal research to improve the quality of investigation.

Other methodological approaches are also needed, either to fill the gaps that cannot be addressed by RCTs or to focus on research questions that could never be addressed by a very structured, scientific approach. Mixed method approaches can complement each other and contribute, overall, to a stronger research design for herbal medicine. Observational studies, for example, can allow pragmatic investigation of areas that are not amenable to RCTs. Moreover, practitioners of modern medicine need to be aware of the research issues and how to judge the quality of research evidence in herbal medicine, so they can provide appropriate advice to patients, and move towards integrating effective therapies into their practice.

Although a number of difficulties and limitations are acknowledged, it has been established that the RCT will continue to be the gold standard for evaluating the efficacy and safety of herbal products. To conclude, it would be helpful if reports of all clinical studies gave more information about the rationale for the choice of research question and related design.

Disclosures

The authors report no conflicts of interest relevant to this research.

References