

Occasional essay

Complementary or alternative medicine: the need for plausibility

L. John Hoffer

Everyone familiar with scientific medicine understands the importance of the randomized controlled trial (RCT), because it is the gold standard for evaluating treatment efficacy. Proof that a treatment is efficacious typically requires at least one *definitive* RCT or several *convincing* ones. It is not sufficient for the RCT to demonstrate a statistically and clinically significant effect; it must also be designed, and its results analyzed, in accordance with rigorous criteria set by clinical trial authorities. There is a trend toward seriously crediting only RCTs with large numbers of participants, and this calls for a complex study design and infrastructure.

It is not always appreciated that such high-cost, definitive RCTs come near the end, not the beginning, of the process of evaluating new therapies. Before the definitive RCT, there is usually a lengthy process of information gathering to rule out toxicity, optimize the parameters of the treatment and determine the clinical importance of its apparent effect.¹ Only sufficiently promising therapies merit the effort and expense of final confirmation (or refutation) with large, definitive RCTs. This preliminary research may be termed “plausibility building.” When biochemical, tissue and animal data point to a mechanism of action or demonstrate the desired biological effect, they thereby confer *biological plausibility*. Clinical data from epidemiological studies, case reports, case series and small, formal open or controlled clinical trials may confer *clinical plausibility*. A therapy is sufficiently scientifically plausible to merit the time and expense of definitive testing if it is *either* biologically *or* clinically plausible.

These considerations are germane to discussions underway in Canada and the United States about the best ways to evaluate complementary or alternative medicine (CAM). Issues surrounding CAM are complex; indeed, even defining CAM can be difficult.² In this article, I use the definition adopted by the US National Library of Medicine and the US National Institutes of Health National Center for Complementary and Alternative Medicine: a CAM therapy is one that is used instead of (“alternative”) or in addition to (“complementary”) the conventional accepted therapy for a condition (www.nlm.nih.gov/nccam/background.htm#c). But what formula determines which therapies are accepted and which ones rejected by conventional medicine? Generally speaking, a therapy joins the canon of conventional, accepted therapies either after its efficacy has been demon-

strated in well-designed clinical trials, or because its biological rationale fits plausibly within the scientific biomedical conceptual framework, even if proof of its efficacy is lacking — it “makes biological sense.” The latter stipulation is necessary, because many conventional therapies are unproven but still not considered CAM. The definition of CAM thus hinges on the notion of scientific plausibility: CAM therapies are not considered conventional medicine because they lack good evidence of clinical efficacy (clinical plausibility) and they lack biological plausibility.

As obvious as it is, this definition can be difficult to apply in practice. The recognition and implementation into medical practice of even well-proven interventions is frequently delayed. Judgements can differ over the level of biological or clinical plausibility in a given case. Alternatively, a therapy may be reasonably biologically plausible but not submitted to definitive clinical testing because it is not sufficiently fashionable or financially rewarding. An additional complication is that, as with homeopathy, the stranger and more biologically implausible a therapy, the higher the bar medical scientists tend to set for crediting evidence supporting its clinical plausibility. Finally, some therapies commonly considered as CAM, such as glucosamine sulfate for osteoarthritis and St. John’s wort for depression, no longer really fit the definition, because well-designed RCTs indicate that they are probably efficacious, but the label continues to stick.

Many CAM therapies (like glucosamine sulfate and St. John’s wort) are easily tested using conventional RCT designs, for they are simple drug or drug-like products with standard clinical indications. Other CAM therapies are far more difficult to evaluate, either because of their complexity or because of the nature of the alternative medical philosophies from which they are derived. Some CAM approaches use definitions of health, diagnosis and disease that differ radically from those used in conventional scientific medicine. Some are more properly considered lifestyle, cultural or spiritual practices whose innate values go beyond the pathophysiological focus of conventional medical science.³ Despite these complexities, most observers advocate the scientific evaluation of CAM, while acknowledging the considerable difficulties involved.^{4,5}

I too believe that most CAM therapies are amenable to evaluation using RCTs. It is common in RCTs of drugs to

use placebos to blind the participants and investigators as to treatment allocation, but this is not always feasible with CAM approaches. Fortunately, it is now widely recognized that well-designed nonblinded RCTs can generate important conclusions.⁶⁻⁹ Problems can emerge when the attempt is made to randomize clinical trial participants into placebo, nontreatment or alternative treatment groups they did not freely choose, because CAM therapies tend to be harmless, very complicated or require active participation.^{5,10} One way to deal with this is cluster randomization, in which the unit of randomization is a hospital ward, medical practice or community.¹¹⁻¹⁵ Thus, while the barriers to the practical testing of CAM are real, they are not insurmountable.

In fact, as important as RCT design can be, I believe there are more fundamental problems to be solved on the way to establishing rational and fair procedures for determining the merits and flaws of CAM: (1) Which of the large and ever-increasing number of candidate therapies should be tested, and in what order of priority? (2) Who will fund and (3) who will carry out the clinical trials? These problems — and the key to their resolution — arise from the defining feature of CAM: its scientific implausibility.

For a therapeutic approach to be considered CAM it must, by definition, be scientifically implausible. But if a treatment is not scientifically plausible, how is one to proceed with its scientific evaluation? Thorough exploration of the potential of any therapeutic approach requires a notion of its target patients and their response and, whenever possible, its mechanism of action. Without this, testing CAM with large RCTs is liable to be as useful as firing cannons at flocks of sparrows.

What scientific peer review committee would — or should — award a competitive grant for a large RCT to test a treatment that is highly implausible? Private funding cannot be counted on, because most CAM therapies are not patented. And what private interest would willingly fund the definitive test of an already profitable CAM therapy? The result might be negative!

Few career medical researchers do more than dabble in CAM in their spare time, and this is hardly surprising. What bright young researcher would choose to devote a scientific career to confirming the inefficacy of implausible treatments? For their part, most CAM practitioners lack both training and sophistication in clinical investigation and the protected time necessary to gain research skills and conduct clinical trials. Most authorities believe that fruitful efforts to evaluate CAM will require partnerships between CAM practitioner-proponents and skilled clinical research investigators. How will they be created?

The correct way to meet these challenges is to put CAM therapies through the same plausibility-building process that conventional therapies undergo before they come to the stage of definitive RCTs. The premise of this strategy is that a gain in plausibility is not proof. (Conversely, lack of proof does not exclude plausibility.^{1,16}) Rather, as plausibility increases, the case for definitive

RCT testing becomes stronger. The purpose of this research is to groom CAM therapies into serious candidates for definitive testing by RCTs.

A candidate CAM therapy need not be biologically plausible to merit testing. The important drug classes were identified by empirical observation, and in-vitro empirical testing continues to be the most common path to new drug discovery. Indeed, an unexpected drug effect is likely to be more important than one that was predicted, for it can lead to new biological insight.

Biological plausibility is helpful, however. A biologically plausible mechanism of action greatly aids the investigator in selecting an appropriate dosing regimen and patient population, and encourages persistence in the face of disappointing results. The hypothesis that L-dopa would benefit patients with Parkinson's disease appeared to be refuted when well-designed double-blind RCTs turned out to be negative, but its biological plausibility persuaded George C. Cotzias to test much higher doses of L-dopa than had previously been used. His landmark paper in the *New England Journal of Medicine* was a noncontrolled case series.¹⁷

Despite its importance, biological plausibility is lacking for most CAM therapies. This does not mean they are ineffective, but it does require their evaluation to proceed solely on the basis of clinical plausibility. Several evaluation methods are known, but until now they have largely been discounted as unscientific. They are nonrandomized or open clinical trials, case studies and case series. Nonrandomized clinical trials can be valuable, as they tend to predict the results of subsequent definitive RCTs with reasonable accuracy.¹⁸ Many case study designs are possible, and the process of data collection and interpretation can be standardized to maximize reliability and validity.¹⁹ An accumulation of informative case histories can be developed into a "best case series."²⁰ The US National Center for Complementary and Alternative Medicine has developed guidelines for preparing a best case series of responses to an alternative cancer therapy (www3.cancer.gov/occam/bestcase.html).

An option available in some situations is the RCT of individual patients. In such a study, the patient acts as his or her own control.^{21,22} This approach (also termed the "n-of-1" study) was formalized at McMaster University, Hamilton, Ont., where it proved so popular that a clinical service was established to facilitate its use in the community. The McMaster group cautioned that a positive n-of-1 experiment does not prove that a treatment is effective for all patients with a given disorder.²¹ But the plausibility-building effect of several n-of-1 studies with coherent, positive results would be undeniable.

More convincing than a best case series is a consecutive case series, because it permits more general conclusions. This would be attempted only after the therapeutic algorithm, the appropriate patient population and the outcome variables of interest were clearly defined. Where would a consecutive case series lead? To the publication of further noncontrolled consecutive case series by independent in-

investigators. Several independent positive consecutive case series would further increase plausibility, as long as negative studies were also made public. White and Ernst²³ argue strongly for the use of formal, noncontrolled clinical trials in CAM research. The proper use of case studies and case series requires as much intellectual rigour as a clinical trial, but is more practical for CAM investigators in the field who typically lack funding, protected time or access to the sophisticated resources available in medical school clinical trial units, and whose patients may not be interested in participating in RCTs.

The first priority in a practical effort to focus CAM research productively should be to create mechanisms that foster research into clinical plausibility building. What are needed are (1) the development of formal standards and guidelines for formal case studies, case series and noncontrolled clinical trials in CAM; (2) the recruitment of expert consultants to advise and instruct CAM practitioners; (3) the creation of educational opportunities for CAM practitioners to develop expertise in research into plausibility building; (4) grant competitions for protocols for research into plausibility building and (5) a central registry to record the results (both positive and negative) of plausibility-building research studies.

Without such a vetting process in place, I believe there is a real danger that public funds earmarked in good faith for CAM therapy research will be dissipated in a variety of ways: in descriptive sociology, in pseudo-CAM projects that are really artfully repackaged mainstream research, and in large, mostly futile RCTs of CAM therapies selected on the basis of advocacy rather than merit.

In summary, the way to prove the efficacy of most CAM therapies is with well-designed RCTs, and there is no reason to believe that clinical trial designs cannot be developed that allow even complex CAM therapies to be evaluated. The procedures involved can be sophisticated, complex and expensive, however, and this confronts investigators with the challenge of identifying which of the myriad of existing and future CAM therapies merit the effort and expense of definitive RCT evaluation. The challenge should be met as it is in conventional drug discovery, through plausibility-building research. Whenever possible, efforts should be made to establish a credible mechanism of action for a candidate CAM therapy, because this will increase its biological plausibility and reduce the risk of false-negative RCT results. When biological plausibility is lacking, clinical plausibility alone must be the basis for determining whether or not to proceed to the costlier phase of definitive RCTs. The creation of a plausibility-building CAM research strategy will require thought, instruction, funding, and collaboration among conventional clinical investigators and CAM advocates. The advantages are many: fairness, low cost and the creation of rules of engagement for CAM evaluation that foster balanced partnerships between CAM advocates and mainstream clinical scientists.

This article has been peer reviewed.

Dr. Hoffer is with the Lady Davis Institute for Medical Research, Sir Mortimer B. Davis Jewish General Hospital, and the Department of Medicine, McGill University, Montréal, Que.

Competing interests: None declared.

Acknowledgements: I thank Dr. John Ruedy, Dr. Simon N. Young and the late Dr. Colin Sharpe for helpful comments. This article is dedicated to Colin Sharpe.

References

1. Antman K, Lagakos S, Drazen J. Designing and funding clinical trials of novel therapies. *N Engl J Med* 2001;344(10):762-3.
2. Spencer JW. Essential issues in complementary/alternative medicine. In: Spencer JW, Jacobs JJ, editors. *Complementary/alternative medicine: an evidence-based approach*. Mosby: St.Louis; 1999. p. 3-36.
3. O'Hara DP. Is there a role for prayer and spirituality in health care? *Med Clin North Am* 2002;86:33-46.
4. Ernst E. The role of complementary and alternative medicine in cancer. *Lancet Oncol* 2000;1:176-180.
5. Antman K, Benson MC, Chabot J, Cobrinik D, Grann VR, Jacobson JS, et al. Complementary and alternative medicine: the role of the cancer center. *J Clin Oncol* 2001;19:55S-60S.
6. Appel LJ, Moore TJ, Obarzanek E, Vollmer WM, Svetkey LP, Sacks FM, et al. A clinical trial of the effects of dietary patterns on blood pressure. DASH Collaborative Research Group. *N Engl J Med* 1997;336:1117-24.
7. Sacks FM, Svetkey LP, Vollmer WM, Appel LJ, Bray GA, Harsha D, et al. Effects on blood pressure of reduced dietary sodium and the dietary approaches to stop hypertension (DASH) diet. DASH-Sodium Collaborative Research Group. *N Engl J Med* 2001;344:3-10.
8. Hrobjartsson A, Gotzsche PC. Is the placebo powerless? An analysis of clinical trials comparing placebo with no treatment. *N Engl J Med* 2001;344:1594-9.
9. Knowler WC, Barrett-Connor E, Fowler WE, Hamman RF, Lachin JM, Walker EA, et al. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *N Engl J Med* 2002;346:393-403.
10. Hart A. Randomized controlled trials: the control group dilemma revisited. *Complement Ther Med* 2001;9:40-4.
11. Byar DP. Some statistical considerations for design of cancer prevention trials. *Prev Med* 1989;18(5):688-99.
12. Inouye SK, Bogardus ST Jr, Charpentier PA, Leo-Summers L, Acampora D, Holford TR, et al. A multicomponent intervention to prevent delirium in hospitalized older patients. *N Engl J Med* 1999;340:669-76.
13. Donohoe ME, Fletton JA, Hook A, Powell R, Robinson I, Stead JW, et al. Improving foot care for people with diabetes mellitus — a randomized controlled trial of an integrated care approach. *Diabet Med* 2000;17:581-7.
14. Walker AE, Campbell MK, Grimshaw JM. A recruitment strategy for cluster randomized trials in secondary care settings. *J Eval Clin Pract* 2000;6:185-92.
15. Gortmaker SL, Peterson K, Wiecha J, Sobol AM, Dixit S, Fox MK, et al. Reducing obesity via a school-based interdisciplinary intervention among youth: Planet Health. *Arch Pediatr Adolesc Med* 1999;153:409-18.
16. Hoffer LJ. Proof versus plausibility: rules of engagement for the struggle to evaluate alternative cancer therapies. *CMAJ* 2001;164(3):351-3.
17. Cotzias GC, Papavasiliou PS, Gellene R. Modification of Parkinsonism — chronic treatment with L-dopa. *N Engl J Med* 1969; 280(7):337-45.
18. Ioannidis JP, Haidich AB, Pappa M, Pantazis N, Kokori SI, Tektonidou MG, et al. Comparison of evidence of treatment effects in randomized and nonrandomized studies. *JAMA* 2001;286:821-30.
19. Lukoff D, Edwards D, Miller M. The case study as a scientific method for researching alternative therapies. *Altern Ther Health Med* 1998;4:44-52.
20. Vanchieri C. Alternative therapies getting notice through Best Case Series Program. *J Natl Cancer Inst* 2000;92:1558-60.
21. Guyatt G, Sackett D, Taylor DW, Chong J, Roberts R, Pugsley S. Determining optimal therapy — randomized trials in individual patients. *N Engl J Med* 1986;314(14):889-92.
22. Guyatt G, Sackett D, Adachi J, Roberts R, Chong J, Rosenbloom D, et al. A clinician's guide for conducting randomized trials in individual patients. *CMAJ* 1988;139:497-503.
23. White A, Ernst E. The case for uncontrolled clinical trials: a starting point for the evidence base for CAM. *Complement Ther Med* 2001;9:111-6.

Correspondence to: Dr. L. John Hoffer, Lady Davis Institute for Medical Research, Sir Mortimer B. Davis Jewish General Hospital, 3755 Côte-Ste-Catherine Rd. W, Montréal QC H3T 1E2; fax 514 340-7502; l.hoffer@mcgill.ca