



Exposing the evidence gap for complementary and alternative medicine to be integrated into science-based medicine

Michael Power¹ • Kevork Hopayian²

¹NHS Clinical Summaries, Sowerby Centre for Health Informatics Newcastle Limited (SCHIN), Newcastle upon Tyne, UK

²School of Medicine, Health Policy and Practice, University of East Anglia, Norwich, UK

Correspondence to: Michael Power. Email: Michael.power@schin.co.uk

DECLARATIONS

Competing interests

None declared

Funding

None

Ethical approval

Not applicable

Guarantor

MP

Contributorship

Both authors contributed equally

Acknowledgements

The authors thank Iain Chalmers, Imogen Evans, Paul Glasziou, Andy Oxman, Franz Porzolt, Merrick Zwarenstein and the reviewers for their helpful comments on earlier draft of this paper. The opinions expressed in this article are those of the authors, and do not reflect those of their employers or their advisors

Summary

When people who advocate integrating conventional science-based medicine with complementary and alternative medicine (CAM) are confronted with the lack of evidence to support CAM they counter by calling for more research, diverting attention to the ‘package of care’ and its non-specific effects, and recommending unblinded ‘pragmatic trials’. We explain why these responses cannot close the evidence gap, and focus on the risk of biased results from open (unblinded) pragmatic trials. These are clinical trials which compare a treatment with ‘usual care’ or no additional care. Their risk of bias has been overlooked because the components of outcome measurements have not been taken into account. The components of an outcome measure are the specific effect of the intervention and non-specific effects such as true placebo effects, cognitive measurement biases, and other effects (which tend to cancel out when similar groups are compared). Negative true placebo effects (‘frustratebo effects’) in the comparison group, and cognitive measurement biases in the comparison group and the experimental group make the non-specific effect look like a benefit for the intervention group. However, the clinical importance of these effects is often dismissed or ignored without justification. The bottom line is that, for results from open pragmatic trials to be trusted, research is required to measure the clinical importance of true placebo effects, cognitive bias effects, and specific effects of treatments.

Introduction

Integrative medicine, the integration of complementary and alternative medicine (CAM) with conventional healthcare, is being increasingly adopted within mainstream health services. IM advocates, faced with a growing body of evidence that questions the effectiveness of CAM therapies

and their underlying theories, have responded vigorously using strategies that call for more research when there is evidence of no clinically important benefit; focus on the ‘package of care’ rather than specific interventions; talk vaguely about ‘non-specific effects’ rather than precisely about placebo effects and bias effects; and call for unblinded pragmatic trials of CAM without

considering the risks of bias. We explain why each of these strategies fails to close the gap between the evidence and the claims of integrative medicine – the integrative medicine evidence gap.

Our focus on problems with research in CAM and integrative medicine is a specific response to attempts intended to bring them into the mainstream. It does not imply that the evidence base in conventional medicine does not also have major problems.^{1,2}

Why ‘more research is needed’ does not close the integrative medicine evidence gap

When research provides good evidence that CAM treatments have no clinically important benefit, advocates of CAM call for more research.³ Superficially, this suggests a scientific and open mind. However, a chasm exists between the even-handed skepticism characteristic of science and the partisan denial characteristic of CAM. The integrative medicine community presumes that CAM is effective because practitioners and patients believe it to be.⁴ For example, rigorous syntheses of a weighty body of evidence could not unearth any benefit from homeopathy⁵ but homeopaths denied the value of scientific evaluation and concocted a theory, no more plausible than that of homeopathy itself, to explain why such evaluations fail.⁶ CAM advocates, denying the evidence of lack of effectiveness, call for even more research despite the expenditure of very large sums on research that fails to provide what they would regard as convenient evidence. For example, in the USA, the \$1.288 billion spent since 1999 by the National Center for Complementary and Alternative Medicine (NCCAM) on research into CAM, has produced no reliable evidence of benefit from any CAM modality.⁷

Scientific skepticism assumes that treatments are ineffective and unsafe until there is good evidence to the contrary. Progress in scientific endeavours such as medicine and engineering, comes from testing predictions derived from hypotheses, and when predictions fail, the falsified hypotheses are rejected and new hypotheses developed for testing – a process termed the hypothetico-deductive model.^{8,9}

Thus, when a body of good evidence indicates that a therapy has no clinically useful specific

effects, scarce resources should not be wasted in conducting more research.

Why focusing on the ‘package of care’ does not close the integrative medicine evidence gap

Advocates of integrative medicine say that it is ‘the effect of the package of care delivered by the therapist or therapists that is of interest rather than the individual components of the treatment package’.¹⁰ So, for example, the needling in acupuncture, the spinal manipulation in chiropractic, and the pillule in homeopathy are not of particular interest because the relevant effects are due to the whole package of care.^{10–13}

Focusing on the ‘package of care’ is rhetorically convenient for two reasons. First, it camouflages the embarrassing (for integrative medicine) fact that well-controlled and adequately blinded studies of CAM interventions fail to find clinically useful specific effects.¹⁴ Secondly, the package of care may seem to have clinically important effects,¹⁵ but, as we will explain when we discuss non-specific effects, things are not always what they seem.

Concentrating on the package of care, although convenient, is misleading and unscientific: it suggests that it does not matter if the intervention has no useful specific effects and that we should not bother to unpack the package to find out which components do what. Because this approach could be used to justify any bogus medicine, drugs must have evidence of specific effects before they are licensed.

While scientific medicine regards it as essential to investigate specific effects, it has long recognized the importance of the package of care and its components have been the subject of much research, including the roles of expectation,¹⁶ and communication,¹⁷ and the therapeutic relationship between doctor and patient.^{4,18}

Mistaking the package of care as unique to CAM and concluding that the package should not be unpacked deflects us from important research endeavours and treatment opportunities. We should identify and exploit those components that are effective and discard those that are ineffective or even harmful.

Why focusing on 'non-specific effects' does not close the integrative medicine evidence gap

Advocates of integrative medicine use the term 'non-specific effect' in preference to 'placebo effect', perhaps because placebo effect has negative associations⁴ and is sometimes used with contempt or disapproval as in 'it's only a placebo', or 'placebo treatments are unethical'.

The terms non-specific effects and placebo effects are often assumed to be equivalent and are often used without explanation. This hinders clear thinking because neither term accurately describes its meaning. Placebo effects are not due to the placebo medicine (or sham treatment) because, by definition, placebos have no specific effects. And, non-specific effects can be quite specific but arise from unspecified causes.

To understand the difference between true placebo effects and non-specific effects we need first to understand what a clinical trial measures. In a clinical trial, an outcome measurement is the net result of four types of effect: (1) the specific effect of the intervention; (2) the non-specific effects that are due to the context and process of delivery of the intervention and how outcomes are measured; (3) random variations in all components of the measurement; and (4) errors such as innocent mistakes or deliberate fraud. Here we can ignore random variation (this evens out in the long run) and errors (they will be discovered eventually).

The true placebo effect is the clinical effect of the package of care that is not due to the specific intervention. The non-specific effect (or perceived placebo effect¹⁹) is the observed effect of the package of care in the placebo group of a clinical trial; as illustrated in Table 1 it is the net result of the true placebo effect plus systematic measurement errors (biases).¹⁵

Some systematic measurement errors are the same in the different arms of a clinical trial and so cancel out when the different arms are compared. These include the natural course of the disease, instrument biases and regression to the mean. Other systematic measurement errors are likely to be different in the different arms of a clinical trial and therefore introduce errors such as cognitive measurement biases into comparisons between different groups.

Cognitive measurement biases are errors of attribution, recollection or reporting that lead to systematic errors in outcome measurements. They may originate in the patient, the therapist or the assessor of outcome measures. Cognitive biases can cause a trial participant to report what they think the trial organizers want/ought to hear, or cause a practitioner to recall mostly favourable outcomes.

Placebo-controlled clinical trials can separate out an intervention's specific effect, because it is the difference between the observed effects in the intervention and placebo groups. However, special studies are required to separate true placebo effects from other non-specific effects. Such studies are seldom performed, but this does not mean that they are unimportant. True placebo effects and cognitive measurement biases *must* be distinguished for two reasons. First, true placebo effects are welcome when they are positive – research into enhancing and exploiting them should thus be encouraged. Second, cognitive measurement biases diminish the reliability of clinical trials; more research is required to estimate their importance and to provide guidance on how to reduce the risk that the results are biased.

Why 'a different research approach is needed' does not close the integrative medicine evidence gap

Integrative medicine proponents argue that blinded, placebo-controlled, randomized trials are unsuitable for studying CAM and that unblinded pragmatic trials are therefore essential.^{3,10,20} This argument is based on a misunderstanding of the distinction between 'explanatory' and 'pragmatic' attitudes in therapeutic trials drawn by Schwartz and Lelouch in 1967.²¹

Explanatory trials and pragmatic trials differ in their purpose and design²² and have valuable but distinct roles. Explanatory clinical trials minimize bias and variation and the intervention is often compared with a placebo. Explanatory trials are best suited to answer the question 'Does the intervention work in ideal circumstances and if so, by how much?' Pragmatic clinical trials are less prescriptive in their design, often permit some

Table 1**Components of non-specific effects in the experimental and comparison groups of a clinical trial**

<i>Experimental group</i>	<i>Comparison group</i>	<i>Individual effects</i>	<i>Net effect</i>
Regression to the mean	Regression to the mean	} These effects are similar in the two groups, and so they cancel each other out	No difference between the two groups
Temporal variation	Temporal variation		
Instrumental measurement biases	Instrumental measurement biases		
True placebo effect		} Positive non-specific effects in the experimental group produce a <i>perceived</i> benefit for the experimental group	Perceived net beneficial effect of the intervention
Positive cognitive bias in the observer			
Positive cognitive bias in the patient		} Negative non-specific effects in the comparison group produce a <i>perceived</i> benefit for the experimental group	
	Negative cognitive bias in the patient		
	Negative cognitive bias in the observer		
	Frustratebo effect		

For brevity, some effects are not shown, for example, random variation (which evens out in the long run) and errors such as innocent mistakes and deliberate fraud, which will be discovered when studies are independently repeated

degree of variation in the delivery of the experimental treatment, and the comparison is often with 'usual care'. Ideally, pragmatic trials are conducted after explanatory trials have provided supporting evidence that the intervention can work in ideal circumstances. Pragmatic trials answer the question 'How well does this intervention, work in usual practice?' Compared with explanatory trials, pragmatic trials provide a package of treatment closer to that given in routine practice. *Sometimes* that means that the therapist or the patient or both cannot be blinded but it is not always so. Blinding of assessors, randomization, and the use of objective assessments of outcomes are all possible, indeed desirable in pragmatic studies.²³

Some CAM proponents have argued that using a placebo predisposes to type II errors (falsely concluding that there is no effect) because the non-specific effects are experienced by both intervention and placebo groups.¹⁰ Far from it. With the non-specific effects controlled for in this way, the specific effects of CAM can be truly explored.

Others mistakenly assume that pragmatic trials are never blinded.¹² Some even see the lack of blinding as an advantage.¹³ This belief is 'justified' with fallacies such as 'In a pragmatic trial, it is not

usually appropriate to use a placebo control and blinding, as these are likely to have a detrimental effect on the trial's ecological validity'.¹² Or, lack of blinding should 'maximize synergy' and 'optimize non-specific effects'.¹² Or, 'clinician and patient biases are not necessarily viewed as detrimental in a pragmatic trial but accepted as part of physicians' and patients' responses to treatment and included in the overall assessment'.²⁴

Failure to recognize the distinction between true placebo effects and observed non-specific effects can lead to serious misinterpretations of the trial results, as we explain in the next section.

Why unblinded pragmatic clinical trials do not close the integrative medicine evidence gap

In an unblinded pragmatic trial, the difference between outcomes in the CAM and usual care groups is commonly misunderstood to provide evidence for beneficial effects of the treatment.^{12,13,20} However, these interpretations overlook the important risks of bias illustrated in Table 1.

The first risk of bias is that there is a negative placebo response in the usual care group. This is especially important when patients are recruited with expectations about the value of a CAM and are then disappointed not to receive it. Negative placebo effects are also called *nocebo* (Latin 'I shall harm') effects, but the term is better used for unwanted adverse effects that arise from taking a placebo.²⁵ We suggest a better term for the negative clinical response elicited by not receiving one's preferred treatment might be *frustrebo* (Latin 'I shall disappoint'). A similar concept, the non-specific effects in trial participants assigned to a less preferred treatment, is 'resentful demoralization'.²⁶ Resentful demoralization effects include *frustrebo* effects and cognitive measurement bias effects. *Frustrebo* effects are biases because they make the difference between the CAM group and usual care group seem like a benefit for the CAM group.

Another source of bias is cognitive measurement bias. This is a particular risk with subjective patient-reported outcome measures of pain and disability. The directions of these cognitive biases would be expected to be the same as the directions of placebo and *frustrebo* effects: positive in the CAM group and negative in the usual care group.

What is the evidence on biases from non-specific effects?

Disappointingly little research has been done to measure the individual contributions of the *frustrebo* effect and cognitive measurement bias in open clinical trials, and we were unable to locate any directly relevant studies. Nevertheless, it is inconceivable that this particular human activity should be free of these biases given their ubiquitous presence elsewhere.²⁷

Indirect evidence on the combination of *frustrebo* and cognitive measurement bias effects ('resentful demoralization') could be provided by 'patient preference trials', or preferably a meta-analysis of such trials. One meta-analysis has been conducted.²⁸ It is difficult to draw general conclusions from the meta-analysis, because the results are dominated by trials in which participants had only mild preferences for the treatments, and therefore probably small *frustrebo* and cognitive bias effects. Trials of CAM

interventions such as acupuncture generally find that almost all people would prefer to be given the test treatment,^{28,29} and thus *frustrebo* and cognitive bias effects would be expected to be larger.

Evidence (albeit indirect) that the potential for cognitive bias is greater when outcomes are self-reported and subjective rather than objective is provided by a large systematic review.³⁰ The study reviewed 146 meta-analyses of 1346 trials and found that for studies with subjective outcomes, lack of blinding was associated with exaggerated estimates of effect (ratio of odds ratios 0.75; 95% confidence interval [CI] 0.61–0.93). In contrast, there was little evidence of bias in trials with objective outcomes (ratio of odds ratios 1.01, 95% CI 0.92–1.10).³⁰ It is not surprising therefore that hard outcomes, such as survival rates, are rarely used in CAM trials.

Conclusions

We have explained the muddled thinking underlying several strategies aimed at closing the integrative medicine evidence gap: to call for more research when there is good evidence of no clinically important benefit; to focus attention on the 'package of care' rather than the specific intervention; to promote clinical trials that measure non-specific effects rather than specific effects; and to promote unblinded pragmatic trials without due concern for the several risks of bias in their results.

We have shown that it is important to distinguish between true placebo effects and non-specific effects. We have explained that non-specific effects provide an illusion of true benefit because they include not only true placebo effects (which are to be welcomed), but also bias effects (which, if they cannot be avoided, should at least be taken into account). And, we have demonstrated that results from unblinded pragmatic clinical trials are plausibly at high risk of several sources of bias.

This muddled thinking has affected funding and clinical policy decisions. For example on the basis of evidence on non-specific effects from unblinded pragmatic trials the German Federal Joint Committee of Physicians and Health Insurance Plans made acupuncture for low back pain and knee pain benefits of German health insurance,³¹ and the National Institute for Health and

Clinical Excellence recommended acupuncture, spinal manipulation and exercise as treatment options for persistent low back pain.¹³

Research programmes are also likely to be adversely affected. For example, a report commissioned by the prestigious King's Fund has supported calls for more unblinded pragmatic trials into CAM without cautioning that results from such trials are at high risk of several biases.²⁰

While more research is needed into the true placebo effect, it is the area of cognitive measurement bias in trials that appears most under-researched. There is scope for routinely assessing the risk of such biases, and frustrebo biases, in pragmatic trials with patient preference allocation.

A re-think is also needed in how research is reported. Guidance on research reporting, such as the CONSORT statement³² and its extensions, might need reviewing to highlight the limitations of open trials with subjective outcome measures, and to recommend reporting participants' preferences and their association with outcomes.

At this time of a drive for efficiency savings in the health services, it is especially important that expenditure should be directed to effective interventions and not be wasted on ineffective ones.

References

- Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. *Lancet* 2009;**374**:86–9
- Laurance J. A nudge in the wrong direction. *BMJ* 2010;**341**:c7248
- Macpherson H, Peters D, Zollman C. Closing the evidence gap in integrative medicine. *BMJ* 2009;**339**:b3335
- Moerman DE. *Meaning, Medicine and the 'Placebo Effect'*. Cambridge: Cambridge University Press, 2002
- Ernst E. A systematic review of systematic reviews of homeopathy. *Br J Clin Pharmacol* 2002;**54**:577–82
- Milgrom LR. Homeopathy and the new fundamentalism: a critique of the critics. *J Altern Complement Med* 2008;**14**:589–94
- National Center for Complementary and Alternative Medicine. *NCCAM Funding: Appropriations History*. Bethesda, MD: NCCAM, 2010. See <http://nccam.nih.gov/about/budget/appropriations.htm> (last checked 22 December 2010)
- Hopayian K. Why medicine still needs a scientific foundation: restating the hypotheticodeductive model – part one. *Br J Gen Pract* 2004;**54**:400–1
- Hopayian K. Why medicine still needs a scientific foundation: restating the hypotheticodeductive model – part two. *Br J Gen Pract* 2004;**54**:402–3
- Paterson C, Dieppe P. Characteristic and incidental (placebo) effects in complex interventions such as acupuncture. *BMJ* 2005;**330**:1202–5
- Fonnebo V, Grimsgaard S, Walach H, *et al.* Researching complementary and alternative treatments — the gatekeepers are not at home. *BMC Med Res Methodol* 2007;**7**:7
- Macpherson H. Pragmatic clinical trials. *Complement Ther Med* 2004;**12**:136–40
- Savigny P, Kuntze S, Watson P, *et al.* *Low Back Pain: early management of persistent non-specific low back pain*. London: National Collaborating Centre for Primary Care and Royal College of General Practitioners, 2009. See <http://www.nice.org.uk/nicemedia/pdf/CG88fullguideline.pdf> (last checked 11 August 2010)
- Madsen MV, Götzsche PC, Hróbjartsson A. Acupuncture treatment for pain: systematic review of randomised clinical trials with acupuncture, placebo acupuncture, and no acupuncture groups. *BMJ* 2009;**338**:a3115
- Ernst E. Dissecting the therapeutic response. *Swiss Med Wkly* 2008;**138**:23–8
- Crow R, Gage H, Hampson S, Hart J, Kimber A, Thomas H. The role of expectancies in the placebo effect and their use in the delivery of health care: a systematic review. *Health Technol Assess* 1999;**3**:1–96
- Rao JK, Anderson LA, Inui TS, Frankel RM. Communication interventions make a difference in conversations between physicians and patients: a systematic review of the evidence. *Med Care* 2007;**45**:340–9
- Di Blasi Z, Harkness E, Ernst E, Georgiou A, Kleijnen J. Influence of context effects on health outcomes: a systematic review. *Lancet* 2001;**357**:757–62
- Ernst E, Resch KL. Concept of true and perceived placebo effects. *BMJ* 1995;**311**:551–3
- Black C. *Assessing complementary practice: building consensus on appropriate research methods*. London: The King's Fund, 2009. See http://www.kingsfund.org.uk/publications/complementary_meds.html (last checked 11 August 2010)
- Schwartz D, Lellouch J. J Explanatory and pragmatic attitudes in therapeutical trials. *J Chronic Dis* 1967;**20**:637–48
- Thorpe KE, Zwarenstein M, Oxman AD, *et al.* A pragmatic-explanatory continuum indicator summary (PRECIS): a tool to help trial designers. *J Clin Epidemiol* 2009;**62**:464–75
- Zwarenstein M, Treweek S, Gagnier JJ, *et al.* Improving the reporting of pragmatic trials: an extension of the CONSORT statement. *BMJ* 2008;**337**:a2390
- Roland M, Torgerson DJ. What are pragmatic trials? *BMJ* 1998;**316**:285
- Barsky AJ, Saintfort R, Rogers MP, Borus JF. Nonspecific medication side effects and the nocebo phenomenon. *JAMA* 2002;**287**:622–7
- Cook TD, Campbell DT. *Quasi-experimentation: design and analysis issues for field settings*. Chicago, IL: Rand McNally, 1979
- Gilovich T. *How We Know What Isn't So: The Fallibility of Human Reason in Everyday Life*. London: Free Press, 2008
- Preference Collaborative Review Group. Patients' preferences within randomised trials: systematic review and patient level meta-analysis. *BMJ* 2008;**31**:a1864
- Haake M, Müller HH, Schade-Brittinger C, *et al.* German Acupuncture Trials (GERAC) for chronic low back pain:

- randomized, multicenter, blinded, parallel-group trial with 3 groups. *Arch Intern Med* 2007;**167**:1892–8. Erratum in: *Arch Intern Med* 2007;**167**:2072
- 30 Wood L, Egger M, Gluud LL, *et al.* Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *BMJ* 2008;**336**:601–5
- 31 O’Connell NE, Wand BM, Goldacre B. Interpretive bias in acupuncture research? A case study. *Eval Health Prof.* 2009;**32**:393–409
- 32 CONSORT Group. *The CONSORT statement for reporting randomized trials*. CONSORT: Transparent Reporting of Trials, 2010. See <http://www.consort-statement.org> (last checked 11 August 2010)