

Research article

A systematic review of the quality of homeopathic clinical trials

Wayne B Jonas *¹, Rachel L Anderson ², Cindy C Crawford ¹ and John S Lyons ²

Address: ¹Samueli Institute for Information Biology and Department of Family Medicine, Uniformed Services University of the Health Sciences, Bethesda, Maryland, USA and ²Department of Psychiatry and Behavioral Sciences, Northwestern University Medical School, Chicago, Illinois, USA

E-mail: Wayne B Jonas* - wjonas@siib.org; Rachel L Anderson - randerson@mail.public-health.uiowa.edu; Cindy C Crawford - ccrawford@siib.org; John S Lyons - jsl329@nwu.edu

*Corresponding author

Published: 31 December 2001

Received: 13 July 2001

BMC Complementary and Alternative Medicine 2001, 1:12

Accepted: 31 December 2001

This article is available from: <http://www.biomedcentral.com/1472-6882/1/12>

© 2001 Jonas et al; licensee BioMed Central Ltd. Verbatim copying and redistribution of this article are permitted in any medium for any non-commercial purpose, provided this notice is preserved along with the article's original URL. For commercial use, contact info@biomedcentral.com

Abstract

Background: While a number of reviews of homeopathic clinical trials have been done, all have used methods dependent on allopathic diagnostic classifications foreign to homeopathic practice. In addition, no review has used established and validated quality criteria allowing direct comparison of the allopathic and homeopathic literature.

Methods: In a systematic review, we compared the quality of clinical-trial research in homeopathy to a sample of research on conventional therapies using a validated and system-neutral approach. All clinical trials on homeopathic treatments with parallel treatment groups published between 1945–1995 in English were selected. All were evaluated with an established set of 33 validity criteria previously validated on a broad range of health interventions across differing medical systems. Criteria covered statistical conclusion, internal, construct and external validity. Reliability of criteria application is greater than 0.95.

Results: 59 studies met the inclusion criteria. Of these, 79% were from peer-reviewed journals, 29% used a placebo control, 51% used random assignment, and 86% failed to consider potentially confounding variables. The main validity problems were in measurement where 96% did not report the proportion of subjects screened, and 64% did not report attrition rate. 17% of subjects dropped out in studies where this was reported. There was practically no replication of or overlap in the conditions studied and most studies were relatively small and done at a single-site. Compared to research on conventional therapies the overall quality of studies in homeopathy was worse and only slightly improved in more recent years.

Conclusions: Clinical homeopathic research is clearly in its infancy with most studies using poor sampling and measurement techniques, few subjects, single sites and no replication. Many of these problems are correctable even within a "holistic" paradigm given sufficient research expertise, support and methods.

Background

The popularity of complementary, alternative and uncon-

ventional medicine is increasing worldwide. Several surveys have estimated that between 30 and 70% of patients

in developed countries use these practices, depending on the population and modality. [1–3] There is, however, a paucity of scientific research on many complementary and alternative medical (CAM) practices. A recent commentary in the Archives of Internal Medicine (vol. 156, pp. 2162–2164) called for more research into one of the most controversial of these practices, homeopathy, to determine its quality and clinical efficacy, putting aside for the moment its implausibility.

Over the past several decades there have been a number of clinical trials on various forms of homeopathic treatment for a host of disorders from stroke and post-operative complications to arthritis and fibromyalgia to influenza and upper respiratory track infections. [4,5] While the treatment and conditions studied vary, they share the common property of representing clinical outcome research on forms of homeopathic medicine. Several independent systematic reviews published to date on this research have all shown a surprising number of positive results, even among those trials that received high quality ratings for randomization, blinding, sample size and other methodological criteria. [4–6] None of these reviews, however, have looked at a broad spectrum of validity criteria and none have compared the quality of homeopathic research to conventional research using methods that were not derived from a conventional clinical diagnostic framework.

A number of authors have addressed the difficulties encountered in experimental studies of homeopathy. [5,7–12] While these issues have been extensively discussed in the homeopathic literature, the issues are not well known in the conventional literature. The main issue relevant to assessment of research quality is that the theoretical and clinical approaches of homeopathy use different diagnostic and therapeutic taxonomies than conventional medicine. Unlike conventional medicine that attempts to isolate cause – effect links between single treatments and specific diseases, homeopathic medicine focuses on the stimulation of broad healing processes that have influences on a variety of conditions and symptom patterns. [8] Current classical homeopathic practitioners use complex pattern recognition procedures (often computer assisted) to initiate treatment, evaluate patient response and adjust therapy. It is within this so-called "holistic" paradigm that researchers continue to attempt a critical examination of the specific effects of homeopathic therapies as they relate to conventional diagnostic categories. Lack of homeopathic expertise by researchers or difficulties in modification of homeopathic practice goals to fit conventional diagnosis-based research designs can cause confusion in the design and execution of clinical trials and in the evaluation and interpretation of results. For example, a study by Shipley on the treatment of osteoarthritis with the ho-

meopathic remedy *Rhus tox* reported no difference in outcome compared to placebo. [13] The method for identifying patients who fit the pattern most likely to respond to *Rhus tox* was crude, however, reflecting poor homeopathic practice. A follow-up study by Fisher did a detailed evaluation of a set of patients with fibromyalgia who fit the detailed pattern of *Rhus tox*. using expert homeopathic methods. A randomized controlled trial in this group showed increased effects of *Rhus tox* over placebo. [14]

These differing frameworks also complicate our ability to comparatively evaluate these systems. [7–11] The lack of support for homeopathic therapies among Western practitioners stems, in part, from the scarce quantity and perceived poor quality of homeopathic clinical research. However, no studies have directly compared the quality of research in homeopathic to conventional medicine using a quality assessment method that is neutral to their differing diagnostic classifications, treatment and outcome strategies. A meta-analysis of homeopathic clinical trials (on which WJ was the corresponding author) evaluated some specific quality parameters of 89 placebo-controlled trials of homeopathy using commonly used methods for addressing conventional medical studies. [5] This study reported that homeopathic clinical trials scored about the same as conventional trials on the Jadad scale of quality assessment. The Jadad scale contains only 5 items, however, and measures simple components of internal validity. [15] No information about external validity, generalizability, statistical validity, effects when compared to conventional treatments, and other important trial quality items were assessed.

When doing this meta-analysis, it became clear to us that using review criteria based on diagnostic classifications developed for assessing conventional medicine did not address the differing classification frameworks between homeopathy and conventional medicine. There were practically no studies that addressed both classification schemes and very few that preselected a uniform set of patients according to homeopathic criteria before testing (as in the Fisher trial described above). This resulted in a heterogeneity problem, that is, we were mixing very different types of conditions and classifications but reviewing them from the perspective of one of these classification systems. This problem was not solvable by using narrow review criteria created for one (conventional) system. What was needed was a set of valid quality criteria that were neutral to these two systems with very different theoretical bases and approaches. We felt that an evaluation system called "Systematic Review" (SR) developed by Light and Pillemer [16] provided such an approach since the review criteria were not depended on diagnostic homogeneity and so address quite disparate frameworks. SR is comprehensive,

objective and validated quality assessment method yet also with broad in applicability. It was originally derived and extensively applied to heterogeneous systems of intervention in the social, educational and behavioral sciences but has since been expanded to the evaluation of medicine – both conventional and complementary. [17–19] It allows for a comprehensive assessment of research quality across a variety of study designs and intervention areas.

In SR, the methodological characteristics critical for attaining reliable and valid research results are evaluated across a variety of studies and study designs. By quantitatively evaluating components of research design, statistical methods, generalizability, bias and related findings, SRs provide an estimate of over 30 quality variables in a study set. Collectively, these criteria are referred to as "threats to validity." In addition to providing a comprehensive assessment of quality the systematic evaluation of scientific methods provides a framework from which future studies can be designed. In addition to its application in the social and behavioral sciences, SRs have been successfully applied to conventional and complementary medical approaches including nursing care [19], religion and mental health research [17], abortion [18], and physiotherapy. Because it provides an approach to the evaluation of disparate systems, we used the SR method to evaluate a sample of homeopathic clinical trials and compare them to the quality of research in conventional studies previously evaluated by the SR method [20]. Our specific aim was to get representative samples of the "best" research in both homeopathy and conventional medicine and objectively evaluate their quality using a "system neutral" approach to research assessment.

Methods

Literature search

A multiple strategy search was undertaken to identify research on homeopathic treatments that used the strategy from the comprehensive review by Linde, et al. [5] Briefly, this involved 1) the review by Kleijnen, et al. published in 1991, [4] which used an extensive search strategy for Medline and Embase up to 1990; 2) a direct search of Medline from 1966 to August 1995 using the full text terms homeop*, homoeop*, and the Mesh terms homoeopathy, homoeopathy, and alternative medicine and screening of all citations found; 3) contacts with homeopathic researchers, institutions reporting on homeopathic research, homeopathic manufacturers and follow-up on suggestions from these contacts; 4) searching several extensive homeopathic and complementary medicine registries including those of the Woodward Foundation (USA); CISCOM (RCCM, London); AMED (British Library); HomInform (Glasgow, Scotland); IDAG (Amersfoort, Netherlands); and CCRH (India) as well as several individual collections; 5) attending homeopathic meetings

and searching the conference proceedings, abstract e-books, and indexes from meetings and homeopathic books; 6) the references of reviews and trials found; 7) and additional on-line searches of Medline and Embase from 1989 to October 1995. We operationally defined homeopathic research as studies involving serially prepared dilutions selected according to commonly used methods, including classical, clinical, and complex approaches including pathological and isopathic and tautopathic remedy selection.

Study selection

We sought to get representative samples of the best research in homeopathy and conventional medicine. There is a large disparity in the amount of research in conventional and homeopathic medicine. In addition, there are recognized journals in conventional medicine publish the best research, while this is not the case in homeopathy. We therefore used different approaches to identify the best articles in these databases as follows. For homeopathy, articles had to meet the following criteria: (1) use of a homeopathic intervention for a clinical condition in a comparative trial; (2) assessment of the outcome of the intervention using an empirical measure of some type; (3) be prospective and involve a parallel control group for comparison with homeopathic treatment; (4) have sufficient information reported for us to score with SR review criteria; (5) be in English. We included criterion (3 – studies in parallel control groups) because there is a vast literature of case and series outcome reports in homeopathy that are not written for purposes of research and so would not represent the "gold standard" for homeopathic research.

Evaluation of study quality

Each article was reviewed using an established coding form previously used and validated for SRs to assess the characteristics and quality of empirical research. [19–21] In addition, we determined whether the primary study outcome (as determined by the author or, if unclear, selected by the primary reviewer) was improved, unchanged or made worse by homeopathic treatment. All articles were reviewed through a contract by a person specifically trained in the SR method, and has experience in applying it to a variety of study sets. This reviewer had no previous experience or opinion about homeopathy and was hired to apply the SR methods to this data set. In addition, the quality of these reviews were checked for reliability and accuracy on two samples by two other authors who have extensive experience with the SR method and were also neutral about homeopathy (RA and JL). All criterion are scored as either "met" if the quality items was present and adequate or "not met" if the quality item was either missing or inadequately met. The reliability of these review criteria was consistently greater than 0.95 between and

Table 1: Variables and Definitions Used in Systematic Review

Variable	Definition
Sample size	Number of subjects who participated in the research at the initiation of the study
Gender	Percent of sample male/female
Age	Average age of sample
Ethnicity	Percent of identified racial or ethnic categories
Response rate	Sample size divided by the total number of people approached for participation
Attrition rate	Number of subjects finishing study divided by sample size
Sampling Frame	How sample was identified—random, systematic, mixed, none, not indicated
Research Design	Prospective, cross-sectional or retrospective design
Statistics	Based on highest level used: descriptive, univariate, or multivariate
Funding Source	Who paid for research
Control Group	Whether or not a control group was used in design
Reliability	Any indication of the reliability of any measure used.
Random assignment	Use of random assignment to groups
Confounding variable	Any measurement or statistical applications that attempt to identify and control for variables that might influence outcomes independent of treatment.
Threats to validity	
Threats to Statistical Conclusion Validity	
Low statistical power	Rated as any groups < 10 or correlations with fewer than 30 pairs.
Violated assumption	Evidence of non-normal distributions with parametric statistics
Fishing/error rate	More than 10 statistical tests without a Bonferroni (or similar) correction.
Reliability of measures	Failure to test, note or reference information on measures
Reliability of treatment	Failure to test or note consistently of the application of treatments
Random irrelevancies	Absence of effort to make measurements reasonably consistent.
Random heterogeneity representative	Absence of effort to ensure that sample is reasonably
Threats to Internal validity	
History	Repeated measures without temporal control
Maturation	Repeated measures without temporal control on developmentally sensitive outcomes.
Testing	Use of measures that are sensitive to the testing process.
Instrumentation	Use of poorly or uncalibrated measures.
Statistical regression	Study of cases selected from extremes without control group
Mortality	More than 30% of sample did not complete study
Interaction with selection	Evidence that selection into groups might interact with history, maturation, or testing.
Ambiguity of cause	An association that can be interpreted in either direction vis-à-vis cause and effect.
Diffusion of treatment	Contact between experimental and control subjects.
Compensatory work to Equalization	Evidence that groups are knowledgeable about design and might Equal things out.
Compensatory rivalry	Evidence that groups are knowledge about design and might compete with other group members.
Resentful demoralization	Evidence that one group feels disadvantaged through group assignment process.
Threats to Construct Validity	
Inadequate explication	Evidence that choice of measurement operations does not represent the construct.
Mono-operation bias	Use of only one question for central outcome
Mono-method bias	Use of only one measurement approach (e.g self report) for central outcome
Hypothesis guessing	Evidence that subjects might attempt to guess what results should be (absence of appropriate blinding procedures).
Evaluation apprehension	Evidence that subjects might become anxious during assessments.
Experimenter bias	Absence of controls to keep invested parties from participating in the measurement procedures.
Confounding constructs	Range restriction in measurement
With level of constructs	
Threats to External Validity	
Interaction of different Treatment	Failure to assess additional treatment received during study.
Interaction of testing And treatment	Evidence that testing might be related to the treatment so that subjects complete tests differently after treatment.
Restricted generalizability	Evidence of a limited measurement approach that may not generalize.
Sample bias	Evidence of a poorly chosen sample that does not represent the population considered.
Interaction of selection And treatment	Evidence that group assignment is influenced by variables related to the likelihood of response to treatment
Interaction of setting And treatment	Evidence that treatment works only in some settings.
Interaction of history And treatment	Evidence that the time during which the study was done may have an impact on findings.

among reviewers. Finally, we checked for trends in quality of the homeopathic literature over time by analyzing studies separately in three groups – those published before 1980 (n = 17), and from 1980–89 (n = 29) and 1990–96 (n = 13).

The quality scores for this set of homeopathic studies was compared with a random sample of 100 articles previously evaluated from the *New England Journal of Medicine* and the *Journal of the American Medical Association*. [20] JAMA and the NEJM are considered the top two journals in conventional medical research regardless of study design. In addition, we had previously evaluated this set of studies and found it useful as a "gold standard" for conventional medicine. This sample had identical inclusion criteria to the homeopathic literature except they were not required to have used a parallel control group for comparison. This aspect of study design is assessed with two internal valid-

ity criteria in the SR method – "selection" and "interaction with selection" – and so these items were not used to compare study sets. All other criteria operate independent of a control group design and so allow for a direct comparison of these quality items between the homeopathic and conventional study sets. Table 1 describes the criteria used in the SR method. These criteria have been extensively validated and found applicable to a variety of types of empirical research independent of study design. [19–21]

Analysis

Analysis consisted of descriptive statistics for both sets of studies and a comparison of outcomes in those studies randomized and blinded to those that were not using chi-square method. In addition, we compared the number and percent of "threats-to-validity" in both homeopathic and conventional data sets both individually and in their subgroups.

Table 2: Characteristics of the Clinical Trials in Homeopathy by Systematic Review

Descriptive Data	Number	Percent
Controlled trials found	186	--
Trials included in study	59	32
Peer-reviewed publications	47	80
Proceedings or theses	12	20
Number of journals	22	
No funding support indicated	50	85
Funded by private foundation	7	12
Subjects		
Gender stated	30	54
Race and ethnicity stated	0	0
United States Subjects	1	2
Sample size < 100	40	68
Design		
Placebo controlled	23	40
Randomized	30	51
Sampling frame not stated	49	83
Failed to consider confounders	51	86
Measurement		
Reliability reported	0	0
Number screened reported	2	3
Drop out rate reported	21	36
Average drop out rate	--	17
Outcomes		
Number	229	100
Descriptive only	62	27
Improvement from treatment	73	32
No effect from treatment	90	39
Worse from treatment	4	2

Results

Study descriptions

A total of 59 empirical studies evaluating the effects of homeopathic interventions on clinical conditions were included. (Table 2) Forty-seven of these (79.7%) were articles published in peer-reviewed journals. Twelve articles (20.3%) were published as proceedings. The studies were published in 22 different journals, although 20 (34%) were from the *British Homoeopathic Journal*. Most studies (n = 50, 84.7%) indicate no funding support, although two (3.4%) were funded by a government agency and seven (11.9%) stated they were funded by a private donor/foundation.

Subjects

Across all studies that reported the gender of subjects, 60% of subjects were female and 40% male. Seven studies (11.9%) were entirely female samples while two studies (3.4%) were entirely male samples. Nearly half of all studies (n = 27; 45.8%) failed to report the gender of subjects. No study indicated the race or ethnicity of its sample. Thirty-nine studies (66.1%) indicated the national origins of the sample with only one study comprising subjects from the United States (1.7%) and 38 studies comprising residents of other countries (64.4%). The average sample size was 2,135; however, this high average results from a single study of 18,640. The median sample size was 59, with a range from 10 to 934 (excluding the outlier). Two thirds of studies (n = 40; 67.8%) had a sample size below 100.

Design

All studies were prospective and used a comparison group. Twenty-three studies (40%) used a placebo control. Thirty studies (50.8%) used random assignment between experimental and control groups while six (10.2%) used a matched control group. Forty-nine studies (83.1%) did not describe the sampling frame used, that is, they did not describe how the patient sample was identified. For example, one study described the method for random sampling, four studies used a systematic sample frame (e.g. consecutive cases), one study used a mixed frame (e.g. random sampling of consecutive cases), and four studies did not use a sample frame (i.e., convenience sampling). Most studies (n = 51, 86.4%) failed to consider potentially confounding variables in their designs such as open patient selection, severity of illness or non-blind measurement. Seven studies explored potential confounds and found none, while only one study explicitly discussed potential confounds in the design. The proportion of studies in which specific threats to the validity of causal statements were identified can be found in Table 3.

Measurement

No study reported the reliability of measures. Almost all studies failed to report response rates, that is, the number

of patients screened compared to the number who entered the study (n = 57, 96.6%). Of the two studies that reported the proportion of subjects approached who participated, the average response rate was 36.5%. Nearly two thirds of studies (n = 38, 64.4%) did not report the attrition or drop-out rate. For those reporting this item, an average of 16.9% of subjects began the experiment but failed to complete it.

Study Findings

A total of 229 outcome variables were identified, an average of about 4 outcomes per study. Of these, 62 (27%) were only descriptive and no attempt was made to determine how they were affected by the homeopathic treatment. Seventy-three variables (32%) were reported to have statistically significant improvement as a result of the treatment. Ninety outcome variables (39%) were unaffected by the treatment, while four variables (2%) were made significantly worse by the treatment.

The probability of a positive outcome was significantly *lower* when random assignment was used than when random assignment was not used (chi square= 11.9, p < .001) a finding often found in study series and confirmed by others in homeopathic trials. [12] However, when the homeopathic treatment was compared to a conventional medical treatment, the probability of a positive outcome was significantly *higher* than when a placebo control was used (chi square = 15.3, p < .0001).

Overall quality

Significant threats to validity occurred in 30–40% of studies in every major category of validity measured. (Figure 1) Especially high were problems with external validity since most studies had small sample sizes and few were multi-centered. Thus, the results of research in homeopathic medicine is less likely to be generalizable beyond study populations than the results of research in conventional medicine. Surprisingly, internal validity quality scores were very similar between homeopathic and conventional studies, indicating that homeopathic research is not appreciably worse than conventional research when it comes to controlling for bias and systematic error. (Figure 1) Compared to conventional studies using these same quality criteria the homeopathic studies had a higher number of validity threats in 21 categories (63.6%), and the same in 7 (21%) categories. Conventional studies scored worse than homeopathic studies in 5 (15%) categories. How homeopathic and conventional studies compare on all validity threats is shown in Table 3. We were unable to do a subset analysis of conventional studies that used only parallel control groups in order to confirm the independence of validity threats from study design in this sample.

Table 3: Percent of studies with specific threats to validity of homeopathic and conventional interventions.

Statistical Conclusion Validity	Homoeopathic	Conventional
Low Statistical Power	6.8	24.0
Violated Assumptions of Statistical Tests	18.6	14.0
Fishing/Error Rate Problem	6.8	24.0
Reliability of Measures	100.0	94.0
Reliability of Treatment Implementation	16.9	0.0
Random Irrelevancies in Experimental Setting	8.5	0.0
Random Heterogeneity of Subjects	57.6	52.0
Internal Validity		
History	3.4	2.0
Maturation	6.8	2.0
Testing	1.7	0.0
Instrumentation	0.0	0.0
Statistical Regression	3.4	0.0
Selection	100.0	70.0
Attrition	76.3	0.0
Interaction with Selection	20.3	2.0
Ambiguity of Direction of Causality	0.0	2.0
Diffusion of Treatment	1.7	0.0
Compensatory Equalization of Treatments	0.0	0.0
Compensatory Rivalry	0.0	0.0
Resentful Demoralization	1.7	0.0
Construct Validity		
Inadequate Pre-operational Explication of Constructs	22.0	4.0
Mono-Operation Bias	37.3	92.0
Mono-Method Bias	66.1	94.0
Evaluation Apprehension	0.0	0.0
Experimenter Expectancies	3.4	0.0
Confounding Constructs with Level of Constructs	1.7	0.0
Interaction of Testing and Treatments	0.0	0.0
Restricted Generalizability Across Constructs	22.0	0.0
External Validity		
Sample Bias	94.9	52.0
Interaction of Selection and Treatment	1.7	0.0
Interaction of Setting and Treatment	74.6	0.0
Interaction of History and Treatment	0.0	0.0

Quality trends

The overall quality scores in homeopathic trials improved somewhat from 1945 to 1995. Cohort group summary scores went from 30% total validity threats in studies

done prior to 1980 (n = 17), to 26% validity threats between 1980–1989 (n = 29) and 24% between 1990 and 1996 (n = 13) (NS).

Discussion

This review of clinical outcome research in homeopathy has several important implications for the field not revealed in other reviews. Despite the absence of funding, it appears that a number of prospective clinical trials using random assignment and control group designs have been accomplished. The results of these studies suggest that, in general, homeopathic treatments may offer some promise, particularly when compared to conventional medical treatments. As expected, the findings are less impressive when a placebo control and randomization are used. This pattern may reflect the placebo influences common to both conventional and homeopathic therapies and the possibility of selection bias in favor of homeopathy for many studies reported. Both of these factors would falsely elevate the number of positive studies reported in homeopathic research. On the other hand, homeopathic physicians spend considerably more time per visit with patients than conventional physicians. [22] The comprehensive interview and long-term follow-up undertaken as part of homeopathic case management increases patient-doctor interaction and likely enhances expectancy and placebo effects. This may make it more difficult to detect and isolate any additional effects from homeopathic preparations, if these occur. Combined with small sample size this might falsely increase the number of negative studies reported. Thus, the effect of the homeopathic clinical process needs to be studied separately from the examination of any specific (non-placebo) effects that might occur from homeopathic drugs. The former should be examined in pragmatic randomized trials that compare homeopathic with standard therapy and the later should be examined using simple laboratory or clinical models that can be easily repeated by independent investigators.

Average % of Studies Not Meeting Validity Criteria

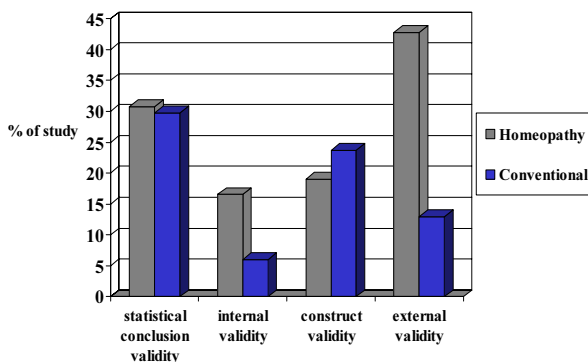


Figure 1
Average % of studies Not Meeting Validity

This study has several limitations. The samples of homeopathic and conventional studies were not identical. While we took all conditions in both study samples, we examined all the English language studies in homeopathy over five decades as opposed to more recent studies only in the conventional literature. In addition, we selected only studies with parallel control groups in homeopathy but took all intervention studies in conventional medicine. We did this in attempt to select a "gold standard" sample of research in each area. While this difference between samples should only have effected two of our 33 quality criteria we cannot be sure that other validity scoring were not effected. To assess the relevance of these sampling differences we examined only the more recent studies in homeopathy. Unfortunately we were unable to do a subset analysis of only conventional randomized, placebo controlled trials in our conventional subset in order to confirm the independence of validity threat from study design as previously shown. It is likely that this would have further improved the quality scores of the conventional subset, however, which would not affect our main conclusions about the general quality of research and threats to validity in these two systems.

A second limitation is that the Systematic Review method used in this evaluation is derived from the social and behavioral sciences and is not commonly used in medical research evaluation. We selected this system, however, precisely because it was broad enough to examine disparate theoretical systems in a more even handed manner than review criteria developed specifically for use in a biomedical framework. While the SR method has been used in conventional and complementary medicine evaluations in the past we realize that the system can only help make general estimates of comparative quality and that individual or summary scores between the two systems should not be viewed as precise. Since limitations in score precision is a problem with all such scoring methods we have tried to highlight only the general and obvious validity threats found.

How do the quality issues in homeopathic research compare with conventional medical research? Overall, the quality of homeopathic research is not as good as research in conventional medicine. The quality of homeopathic trials has only slightly improved over time. Homeopathic studies scored worse than conventional studies on validity threats in over 60% of the items rated compared to 15% of the items in which convention studies scored worse. Most of these errors were on items that would effect the generalizability of the results to other settings rather than on failure to control for systematic errors within studies (bias). Several quality issues are similar between homeopathic and conventional studies. Notably, both groups of research consistently fail to report on the reliability of

measures used. Unreliable outcome measures can markedly distort the accuracy and value of statistical analysis and interpretation of the results. For example, one otherwise very well done study reported on the rates of tonsillectomy under classical homeopathic versus placebo treatment as a main outcome even though the reliability of clinical decisions for obtaining tonsillectomy are extremely poor, thus, making this comparison meaningless. [23]

Currently, the greatest weaknesses in homeopathic research are the variety of unreplicated studies and the small sample sizes. Although 59 studies were included, there was virtually no replication and practically no overlap in the conditions studied. Two thirds of the studies had total sample sizes below 100 and many did not do power calculations to determine the required subject numbers. In order to build credibility within the medical research field, multiple replications and/or extensions using the same or similar approaches to treat the same or similar medical conditions are necessary. Similarly, multi-site research with larger sample sizes is essential to improve the confidence and generalizability of findings. As in all of medicine, success in a treatment at one site often does not always translate into success at other locations with different practitioners and patients. A second major flaw of homeopathic studies in many have a high "attrition rate." This means that the data analyzed and reported on is often lower than the entry sample size due to all causes, such as dropouts, lost or incorrectly filled out forms, etc. A high attrition rate can be a major threat to validity in a study. In addition, homeopathic studies often demonstrated an interaction of treatment setting and treatment indicating that non-specific effects in the environment more often contributed to the outcome than in conventional research.

Finally, several fairly simple improvements in the reporting of research findings can help to heighten the quality of these studies. Most importantly, an outlined discussion of study limitations and potential confounders is important for a balanced interpretation of results and for the advancement of existing research. Similarly, a detailed description of sample characteristics, including gender, ethnicity, and attrition rates is necessary for determining generalizability. By precisely delineating study and sample characteristics, research findings can be stated more definitively. The absence of an appropriate framework from which to examine homeopathic remedies is one but not the sole challenge to quality research in this alternative practice. Critical systematic review of these studies identifies a number of areas for improvement. Quality guidelines for conducting research in homeopathy are published. [24] More critical peer review of homeopathic studies that pay attention to these and other guidelines

can advance quality research in this area. Such criteria need to be carefully addressed by investigators if the quality, credibility and usefulness of research in homeopathy is to improve.

Competing interests

None declared.

Acknowledgements

The authors would like to thank Michelle Shasha for participating in the review of the homeopathic clinical trial articles. This project has been partially supported by a grant, the Standard Homeopathic Foundation and NIH, NC-CAM grant, #AT00178-01.

References

- Eisenberg DM, Davis RB, Ettner S, Appel S, Wilkey S, Van Rompay M, Kessler RC: **Trends in alternative medicine use in the United States 1990-1997: Results of a follow-up national survey.** *JAMA* 1998, **280**:1569-1575
- MacLennan AH, Wilson DH, Taylor AW: **Prevalence and cost of alternative medicine in Australia.** *Lancet* 1996, **347**:569-573
- Fisher P, Ward A: **Complementary medicine in Europe.** *Br Med J* 1994, **309**:107-111
- Kleijnen J, Knipschild P, Riet ter G: **Clinical trials of homeopathy.** *Br Med J* 1991, **302**:316-23
- Linde K, Clausius N, Ramirez G, Melchart D, Eitel F, Hedges LV, Jonas WB: **Are the clinical effects of homeopathy all placebo effects? A meta-analysis of randomized, placebo controlled trials.** *Lancet* 1997, **350**:834-843
- Cucherat M, Haugh MC, Gooch M, Boissel JP: **Evidence of clinical efficacy of homeopathy.** *Eur J Clin Pharmacol* 2000, **56**:27-33
- Heron J: **Critique of conventional research methodology.** *Comp Med Res* 1986, **1**:10-22
- Hornung J, Linde K: **Guidelines for the exact description of the preparation and mode of application of serial dilutions and potencies on ultra low dose effects and homeopathic research - a proposal.** *Berlin Journal of Research in Homeopathy* 1991, **1**,2:121-123
- Patel MS: **Evaluation of holistic medicine.** *Soc Sci Med* 1987, **24**:169-75
- Fisher S, Greenberg R: **How sound is the double-blind design for evaluating psychiatric drugs?** *J Nerv Mental Dis* 1993, **181**:345-350
- Linde K, Jonas WB: **Evaluating complementary and alternative medicine: The balance of rigor and relevance.** In: *Essentials of Complementary and Alternative Medicine* (Edited by Jonas WB, Levin JS) Philadelphia, PA: Lipincott Williams & Wilkins; 1999:57-71
- Linde K, Scholz M, Ramirez G, Clausius N, Melchart D, Jonas WB: **Impact of study quality on outcome in placebo-controlled trials of homeopathy.** *J Clin Epidemiol* 1999, **52**:631-636
- Shipley M, Berry H, Broster G, Jenkins M, Clover A, Williams I: **Controlled trial of homeopathic treatment of osteoarthritis.** *Lancet* 1983, **i**:97-98
- Fisher P, Greenwood A, Huskisson EC, Turner P, Belon P: **Effect of homeopathic treatment on fibrositis (primary fibromyalgia).** *Br Med J* 1989, **299**:365-366
- Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, McQuay HJ: **Assessing the quality of reports of randomized clinical trials: is blinding necessary?** *Controlled Clin Trials* 1996, **17**:1-12
- Light R, Pillemer D: *Summing up: The Science of Reviewing Research.* Cambridge, MA: Harvard University Press; 1984
- Larson D, Sherrill K, Lyons J, Craig FC, Thielman SB, Greenwald MA, Larson SS: **Associations between dimensions of religious commitment and mental health found in the American Journal of Psychiatry and the Archives of General Psychiatry: 1978-1989.** *Amer J Psychiatry* 1992, **149**:557-559
- Lyons J, Larson D, Huckeba W, Mueller C: **Research on the psychosocial sequelae of abortion. 1966 to 1985.** In: *Values and Public Policy, Family Research Council* (Edited by Reiger G); 1988
- Larson D, Lyons J, Hohmann A, Beardsley R, Huckeba W, Rabins P, Lebowitz B: **A systematic review of nursing home research in**

three psychiatric journals: 1966 to 1985. *Int J Geriatric Psychiatry* 1989, **4**:129-134

20. Lyons J, Anderson R, Brule P, Ellison S, Walker S: **Systematic review of the peer-reviewed scientific literature on family, marital, and community variables across twenty fields of inquiry.** *Family Health: From Data to Policy* 1993:43-60
21. Lyons J, Larson D, Bareta J, Liu I, Anderson R, Sparks C: **A systematic analysis of the quantity of AIDS publications and the quality of research methods in three general medical journals.** *Evaluation and Program Planning* 1990, **13**:73-77
22. Jacobs J, Chapman EH, Crothers D: **Patient characteristics and practice patterns of physicians using homeopathy.** *Arch Fam Med* 1998, **7**:537-540
23. de Lange de Klerk E, Blommers J, Kuik DJ, Bezemer PD, Feenstra L: **Effects of homeopathic medicines on daily burden of symptoms in children with recurrent upper respiratory tract infections.** *Br Med J* 1994, **309**:1329-1332
24. Linde K, Melchart D, Jonas WB, Hornung J: **Ways to enhance the quality of clinical and laboratory studies in homeopathy.** *Br Homoeopath J* 1994, **83**:3-7

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMedcentral will be the most significant development for disseminating the results of biomedical research in our lifetime."

Paul Nurse, Director-General, Imperial Cancer Research Fund

Publish with **BMC** and your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours - you keep the copyright



BioMedcentral.com

Submit your manuscript here:

<http://www.biomedcentral.com/manuscript/>

editorial@biomedcentral.com