Are the clinical effects of homoeopathy placebo effects? Comparative study of placebo-controlled trials of homoeopathy and allopathy

Aijing Shang, Karin Huwiler-Müntener, Linda Nartey, Peter Juni; et al

The Lancet; Aug 27-Sep 2, 2005; 366, 9487; Research Library Core pg. 726

Summary

Background Homoeopathy is widely used, but specific effects of homoeopathic remedies seem implausible. Bias in the conduct and reporting of trials is a possible explanation for positive findings of trials of both homoeopathy and conventional medicine. We analysed trials of homoeopathy and conventional medicine and estimated treatment effects in trials least likely to be affected by bias.

Methods Placebo-controlled trials of homoeopathy were identified by a comprehensive literature search, which covered 19 electronic databases, reference lists of relevant papers, and contacts with experts. Trials in conventional medicine matched to homoeopathy trials for disorder and type of outcome were randomly selected from the Cochrane Controlled Trials Register (issue 1, 2003). Data were extracted in duplicate and outcomes coded so that odds ratios below 1 indicated benefit. Trials described as double-blind, with adequate randomisation, were assumed to be of higher methodological quality. Bias effects were examined in funnel plots and meta-regression models.

Findings 110 homoeopathy trials and 110 matched conventional-medicine trials were analysed. The median study size was 65 participants (range ten to 1573). 21 homoeopathy trials (19%) and nine (8%) conventional-medicine trials were of higher quality. In both groups, smaller trials and those of lower quality showed more beneficial treatment effects than larger and higher-quality trials. When the analysis was restricted to large trials of higher quality, the odds ratio was 0.88 (95% CI 0.65–1.19) for homoeopathy (eight trials) and 0.58 (0.39–0.85) for conventional medicine (six trials).

Interpretation Biases are present in placebo-controlled trials of both homoeopathy and conventional medicine. When account was taken for these biases in the analysis, there was weak evidence for a specific effect of homoeopathic remedies, but strong evidence for specific effects of conventional interventions. This finding is compatible with the notion that the clinical effects of homoeopathy are placebo effects.

Introduction

Homoeopathy is a widely used but controversial complementary or alternative therapy.\(^1\)\(^2\) The basic premise is that like is cured by like (\emph{similia similibus curantur})—diseases can be treated by substances that produce the same signs and symptoms in a healthy individual.\(^3\) The preparation of remedies involves serial dilution, commonly to the extent that no molecules of the original substance remain, and vigorous shaking between dilutions (potentisation). During this process information is thought to be transferred from the diluted substance to the solvent,\(^4\) which in the light of current knowledge seems implausible. Many people therefore assume that any effects of homoeopathy must be non-specific placebo effects.\(^5\)

Bias in the conduct and reporting of trials is a possible explanation for positive findings of placebo-controlled trials of both homoeopathy and allopathy (conventional medicine).\(^6\)\(^7\) Publication bias is defined as the preferential and more rapid publication of trials with statistically significant and beneficial results than of trials without significant results.\(^8\) The low methodological quality of many trials is another important source of bias.\(^9\) These biases are more likely to affect small than large studies; the smaller a study, the larger the treatment effect necessary for the results to be statistically significant, whereas large studies are more likely to be of high methodological quality and published even if their results are negative. We examined the effects of homoeopathy and conventional medicine observed in matched pairs of placebo-controlled trials, assessed trial quality and the probability of publication and related biases, and estimated results of large trials least affected by such biases.

Methods

Literature search and data sources

We updated a previous comprehensive search for placebo-controlled trials of homoeopathy, which covered publications up to August, 1995.\(^10\) We searched 19 electronic databases, including specialised homoeopathic and complimentary-medicine registries, covering the period from 1995 to January, 2003: MEDLINE, Pre-MEDLINE, EMBASE, DARE, CCTR, CDSR, CINAHL, AMED, MANTIS, Tostline, PASCAL, BIOL, Science Citation Index, CISCOM, British Homoeopathic Library, the Homoeopathy Abstract page, HornInform Homoeopathic library, NCCAM, and...
SIGLE. The search terms in MEDLINE were (homeop* OR homoeop* OR homoeopathy (MeSH)) AND (placebo* OR placebo (MeSH) OR placebo effect (MeSH) OR sham). Search terms for the other databases were much the same. We also checked the reference lists of relevant papers, including reviews and meta-analyses of homeopathic interventions, and contacted experts in the specialty. There were no language restrictions.

We searched the Cochrane Controlled Trials Register to identify placebo-controlled trials of conventional medicine. This bibliographic database of controlled trials is maintained by the Cochrane Collaboration. As part of an international effort to search systematically healthcare journals worldwide and other sources of information, the collaboration has combined results of electronic searches and searches by hand to create a comprehensive database of trials. We searched issue 1, 2003, of the Cochrane Controlled Trials Register, which included 333,809 bibliographic references.

Study selection
We defined inclusion and exclusion criteria a priori and applied the same criteria to trials of homeopathy and of conventional medicine. Inclusion criteria were: that the trial was controlled and of treatments or preventive measures with clinical outcomes; that it had a parallel-group design with placebo control; that there was random or quasi-random assignment to treatment and placebo groups; and that a written report (eg, journal publication, abstract, thesis, conference proceeding, unpublished report, book chapter, monograph) was available with sufficient data to allow the calculation of odds ratios. We excluded trials of homeopathic “proving” in which remedies are given to healthy individuals to assess their effects, cross-over trials, and N-of-1 trials.

Procedures
We used prespecified criteria to identify outcomes for inclusion in the analyses. The first choice was the main outcome measure, defined as the outcome used for sample-size calculations. If no main outcome was specified, we selected other outcomes, in the order: patients’ overall assessment of improvement; physicians’ overall assessment of improvement; and the clinically most relevant other outcome measure (for example, the occurrence or duration of an illness). Outcomes were selected randomly if several were judged equally relevant. For each homeopathy trial, we identified matching trials of conventional medicine that enrolled patients with similar disorders and assessed similar outcomes. We used computer-generated random numbers to select one from several eligible trials of conventional medicine. Outcomes were selected and trials matched without knowledge of trial results.

We used a piloted data-extraction sheet, which covered descriptive information on the trial and study population, intervention, outcome measures, and trial quality. Data were extracted independently by two observers, and discrepancies were resolved by consensus.

Homeopathic interventions were defined as classical, clinical, or complex homeopathy, or as isopathy. Classical homeopathy was defined as comprehensive homeopathic history-taking, followed by the prescription of a single individualised remedy, possibly with subsequent change of remedy in response to changing symptoms. If no comprehensive homeopathic history was taken and all patients received a single, identical remedy, interventions were classified as isopathy if the agent that was judged to be the cause of the disorder was used (for example, pollen in pollinosis). Indications for treatment were classified as acute or chronic or primary prevention or prophylaxis (interventions with the intention of

<table>
<thead>
<tr>
<th>Clinical topic</th>
<th>Number of trial pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Respiratory tract infections</td>
<td>21 (19%)</td>
</tr>
<tr>
<td>Pollinosis and asthma</td>
<td>10 (15%)</td>
</tr>
<tr>
<td>Gynaecology and obstetrics</td>
<td>14 (13%)</td>
</tr>
<tr>
<td>Surgery and anaesthetics</td>
<td>12 (11%)</td>
</tr>
<tr>
<td>Gastroenterology</td>
<td>12 (11%)</td>
</tr>
<tr>
<td>Musculoskeletal disorders</td>
<td>11 (10%)</td>
</tr>
<tr>
<td>Neurology</td>
<td>10 (9%)</td>
</tr>
<tr>
<td>Other</td>
<td>14 (13%)</td>
</tr>
</tbody>
</table>

Table 1: Distribution of pairs of placebo-controlled trials by clinical topic
preventing the occurrence of a disorder or complication). The duration of follow-up was measured in weeks from the start of the treatment to the assessment of outcomes. Assessment of study quality focused on three key domains of internal validity: randomisation (generation of allocation sequence and concealment of allocation), masking (of patients, therapists, and outcome assessors), and data analysis (by intention to treat or otherwise). Random-number tables, computer-generated random numbers, minimisation, coin-tossing, card-shuffling, and lot-drawing were classified as adequate methods for the generation of the allocation sequence. Sealed, opaque, sequentially numbered assignment envelopes, central randomisation, independently prepared and coded drug packs of identical appearance, and on-site computerised randomisation systems were classified as adequate methods of allocation concealment. Analysis by intention to treat was assumed if the reported number of participants randomised and the number analysed were identical. Descriptions of other methods were coded either as inadequate or unclear, depending on the amount of detail provided. Trials described as double-blind, with adequate methods for the generation of allocation sequence and adequate concealment of allocation, were classified as of higher methodological quality.

Graphical and statistical analysis
We expressed results on the odds ratio scale and used the method described by Hasselblad and Hedges to convert differences in continuous outcomes to odds ratios. We recoded outcomes if necessary, so that odds ratios below 1.0 indicated a beneficial effect of treatment in all cases. We used descriptive analyses to compare characteristics of homeopathy and conventional-medicine trials. We examined heterogeneity between trials with standard χ² tests and calculated I² statistics, which measure the proportion of variation in treatment effect estimates due to between-study heterogeneity.

We investigated the association between study size and trial results in funnel plots, by plotting odds ratios on the horizontal axis (on a logarithmic scale) against their SE on the vertical axis. The extent to which study-level variables were associated with log odds ratios was examined by fitting of univariable and multivariable meta-regression models. The following variables were considered: SE of log odds ratio, language of publication, indexing of the publication in MEDLINE, trial quality (masking, generation of allocation sequence, concealment of allocation, intention-to-treat analysis), duration of follow-up, and clinical topic. For homeopathy trials, we also examined whether effects varied between types of homeopathy and types of indications (acute, chronic, primary prevention, or prophylaxis).

We combined treatment effects from larger trials of higher quality by use of standard random-effects meta-analysis and used meta-regression analysis to predict treatment effects in trials as large as the largest trials included in the study. Trials with SE in the lowest quartile were defined as larger trials. Results are given as odds ratios, ratios of odds ratios, or asymmetry coefficients with 95% CI. Ratios of odds ratios of less than 1.0 correspond to a smaller odds ratio for trials with the characteristic and hence a larger apparent benefit of the intervention. Funnel-plot asymmetry was measured by the asymmetry coefficient: the ratio of odds ratios per unit increase in SE of log odds ratio. All analyses were done in Stata version 8.2.

Role of the funding source
The funding sources had no role in the study design; collection, analysis, or interpretation of data; or the writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit the paper for publication.

Results
We identified 165 potentially eligible reports of placebo-controlled trials of homoeopathy and excluded 60 reports. The commonest reasons for exclusion were insufficient information (precluding the calculation of odds ratios), ineligible study design, multiple publication, and inability to identify a matching trial of conventional medicine (figure 1). We included 105 publications that reported on a total of 110 independent trials of homoeopathy (webappendix 1) and 110 publications of 110 matched trials of conventional medicine (webappendix 2).

<table>
<thead>
<tr>
<th>Type of publication</th>
<th>Homeopathy trials (n=110)</th>
<th>Conventional-medicine trials (n=110)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample size</td>
<td>65 (10-1573)</td>
<td>65 (12-1367)</td>
</tr>
<tr>
<td>Median (SD)</td>
<td>117 (211)</td>
<td>133 (226)</td>
</tr>
<tr>
<td>Type of outcome</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall assessment of response</td>
<td>54 (49%)</td>
<td>49 (45%)</td>
</tr>
<tr>
<td>Occurrence or duration of disorder</td>
<td>26 (24%)</td>
<td>26 (24%)</td>
</tr>
<tr>
<td>Assessment of symptoms</td>
<td>21 (19%)</td>
<td>26 (24%)</td>
</tr>
<tr>
<td>Measurement of function or state</td>
<td>6 (5%)</td>
<td>6 (5%)</td>
</tr>
<tr>
<td>Assessment of clinical signs</td>
<td>3 (3%)</td>
<td>3 (3%)</td>
</tr>
<tr>
<td>Trial quality</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Described as double-blind</td>
<td>101 (92%)</td>
<td>96 (87%)</td>
</tr>
<tr>
<td>Adequate generation of allocation sequence</td>
<td>27 (25%)</td>
<td>30 (27%)</td>
</tr>
<tr>
<td>Adequate concealment of allocation</td>
<td>49 (46%)</td>
<td>21 (19%)</td>
</tr>
<tr>
<td>Analysis by intention to treat</td>
<td>33 (30%)</td>
<td>40 (36%)</td>
</tr>
<tr>
<td>Higher quality*</td>
<td>21 (19%)</td>
<td>9 (8%)</td>
</tr>
</tbody>
</table>

*Trials described as double-blind, with adequate generation of allocation sequence and adequate concealment of allocation.

Table 2: Characteristics of placebo-controlled trials of homeopathy and conventional medicine.
The clinical topics studied in pairs of trials ranged from respiratory infections to surgery and anaesthesia (table 1). The outcomes studied were closely matched; overall assessments of response were analysed in 49% of homoeopathy trials and 45% of trials of conventional medicine (table 2). More detailed information on outcomes is given in the webtable. The average study size was similar for the two groups, with a median of around 65 participants. Overall, study size ranged from ten to 1573 participants. Among homoeopathy trials 48 (44%) concerned clinical homoeopathy, 35 (32%) complex homoeopathy, 18 (16%) classical homoeopathy, and eight (7%) isopathy. For the remaining trial, the nature of the homoeopathic intervention was unclear. 101 (92%) of the conventional-medicine trials investigated drugs, eight (7%) immunotherapy, and one a vaccine. The drugs most frequently tested were non-steroidal anti-inflammatory agents (11 trials), anti-allergy drugs (11 trials), virostatic drugs (11 trials), and antibiotics (seven trials).

53% of homoeopathy trials were published in English compared with 83% of trials in conventional medicine. 50 homoeopathy trials were published in German or French. The two groups of trials also differed in the proportion published in MEDLINE-indexed journals. The two groups had similar methodological quality in terms of masking, generation of allocation sequence, and analysis according to intention to treat, but a higher proportion of homoeopathy trials reported adequate concealment of patients’ allocation. 21 (19%) homoeopathy trials and nine (8%) conventional-medicine trials were of higher quality (table 2).

Most odds ratios indicated a beneficial effect of the intervention (figure 2). SE ranged from 0·01 to 1·65 for homoeopathy trials and 0·13 to 1·52 for conventional-medicine trials. Heterogeneity of trial results was less pronounced for homoeopathy (heterogeneity $\chi^2=309$, df 109, $p<0·0001$) than for conventional medicine (heterogeneity $\chi^2=481$, df 109, $p<0·0001$). This difference is unlikely to be due to chance ($p=0·011$ by $F$ test). The proportion of total variation in the estimates of treatment effects due to between-study heterogeneity ($I^2$) was 65% for homoeopathy and 77% for conventional medicine.

Funnel plots were asymmetrical, with smaller trials (larger SE) in the lower part of the plot showing more beneficial treatment effects than larger trials (smaller SE, figure 2). In meta-regression models, the association between SE and treatment effects was similar for trials of homoeopathy and conventional medicine: the respective asymmetry coefficients were 0·17 (95% CI 0·10–0·32) and 0·21 (0·11–0·40). Therefore, with each unit increase in the SE, the odds ratio decreased by a factor of 0·17 for homoeopathy and 0·21 for conventional medicine (table 3).

Other sources of heterogeneity between homoeopathy trials included the language of publication (more beneficial effects in trials published in languages other than English), indexing in MEDLINE (more beneficial effects in trials not indexed in MEDLINE), and indicators of trial quality (more beneficial effects in trials of lower quality). The effects of these variables were generally similar for conventional-medicine trials but did not reach statistical significance (table 3). There was little evidence that treatment effects varied according to duration of follow-up ($p=0·862$ for homoeopathy, $p=0·594$ for conventional medicine) or clinical topic ($p=0·660$ for homoeopathy, $p=0·360$ for conventional medicine) or that effects differed between different types of homoeopathy ($p=0·636$) or type of indication ($p=0·487$). In multivariable analyses, the SE of the log odds ratio (asymmetry coefficient) was the dominant variable in both groups. Coefficients of other
variables, including study quality, were attenuated and became non-significant.

When the analysis was restricted to the larger trials of higher reported methodological quality, the odds ratio from random-effects meta-analysis was 0.88 (0.65-1.19) based on eight trials of homeopathy and 0.58 (0.39-0.85) based on six trials of conventional medicine. Similarly, for prediction of treatment effects in trials as large as the largest trials, the odds ratio was 0.96 (0.73-1.25) for homeopathy and 0.67 (0.48-0.91) for conventional medicine.

**Discussion**

We compared the effects of homeopathy and conventional medicine that are seen in placebo-controlled trials, examined the presence of bias resulting from inadequate methods and selective publication, and estimated results in trials least affected by these biases. We assumed that the effects observed in placebo-controlled trials of homeopathy could be explained by a combination of methodological deficiencies and biased reporting. Conversely, we postulated that the same biases could not explain the effects observed in comparable placebo-controlled trials of conventional medicine. Our results confirm these hypotheses: when analyses were restricted to large trials of higher quality there was no convincing evidence that homeopathy was superior to placebo, whereas for conventional medicine an important effect remained. Our results thus provide support for the hypothesis that the clinical effects of homeopathy, but not those of conventional medicine, are unspecific placebo or context effects.

In 1991, Kleijnen and colleagues argued that there is no reason to believe that compared with homeopathy "the influence of publication bias, data massage, bad methodology, and so on is much less in conventional medicine". Indeed, we found that trials of homeopathy tended to be of higher methodological quality than conventional-medicine trials, although most trials of either type of medicine were of low or uncertain quality. In both groups, smaller trials and those of lower quality showed more beneficial treatment effects than larger trials and those of higher quality. Between-trial heterogeneity was less pronounced among homeopathy trials. This finding might be expected if heterogeneity between homeopathy trials is essentially due to biased reporting and conduct of trials, whereas in the conventional-medicine sample treatment effects represented an additional relevant source of heterogeneity. When we discussed results with practitioners of homeopathy, they contended that classical homeopathy and homeopathic treatment of chronic disorders, in trials with longer follow-up, would yield specific effects. We addressed these points in additional analyses but found no strong evidence in support of these hypotheses.

This study directly compared the presence of biases and their influence on effect estimates in homeopathy and conventional-medicine trials. Identical definitions were used, and data were abstracted independently by two observers. The search of homeopathic publications was comprehensive, and we are confident that we identified a near-complete set of published placebo-controlled trials of homeopathy. The identification of unpublished studies is notoriously difficult, and we probably missed some of these trials. Conventional-medicine trials were randomly selected from the largest existing database of clinical trials (the Cochrane Controlled Trials Register) and were carefully matched to homeopathy trials for clinical subject and type of outcome.

Different sources of bias are difficult to disentangle. The methodological quality of randomised trials cannot be reliably assessed from published articles because reporting on important features of the methods is incomplete in many cases. Indeed, deficiencies in methods of smaller trials that were either not reported or not assessed by us could also have contributed to the asymmetrical shape of the funnel plot. We have argued elsewhere that the funnel plot should be seen not only as a means of detecting publication bias, but also as a generic tool for examination of small-study effects—the tendency for the smaller studies to show larger treatment effects. If reporting is inadequate, study size can be a more precise measure of trial quality than formal assessments of trial quality. We addressed this possibility by modelling the effects expected in trials as
large as the largest trial included in our study; again, we found little evidence for an effect of homoeopathy but stronger evidence for conventional medicine. Another limitation of our study is the exclusive focus on the beneficial effects of homoeopathy and conventional medicine, rather than on both benefits and risks. However, the trials included in the study were small and lacked the power to reveal infrequent but important adverse effects. Furthermore, reporting on adverse effects is inadequate even in larger trials. A comprehensive and valid assessment of adverse effects would probably not have been possible within the framework of this study.

A previous review, which did not include a meta-analysis, also found that many trials of homoeopathy show beneficial effects but are of low methodological quality. A meta-analysis by Linde and co-workers was based on an extensive literature search, which we updated for our study, but it did not include trials of conventional medicine. These researchers concluded that their results were "not compatible with the hypothesis that the clinical effects of homoeopathy are completely due to placebo". However, in a subsequent, more detailed analysis of the same data, they observed that more rigorous trials yielded smaller effect sizes and that their meta-analysis probably "at least overestimated the effects of homoeopathic treatments". In a separate study, the same group observed that many trials in complementary medicine have important methodological weaknesses. Finally, a study of 23 trials of homoeopathy that were considered to be of high methodological quality found that the few trials that used objective endpoints were all negative.

Our study has implications beyond the question of whether homoeopathic remedies have specific effects. First, an important point to keep in mind is that most systematic reviews and meta-analyses are based on relatively few trials. Simulation studies have shown that detection of bias is difficult when meta-analyses are based on a small number of trials. For example, for the eight trials of homoeopathic remedies in acute infections of the upper respiratory tract that were included in our sample, the pooled effect indicated a substantial beneficial effect (odds ratio 0.26–0.50) and there was neither convincing evidence of funnel-plot asymmetry nor evidence that the effect differed between the trial classified as of higher reported quality and the remaining trials. Such sensitivity analyses might suggest that there is robust evidence that the treatment under investigation works. However, the biases that are prevalent in these publications, as shown by our study, might promote the conclusion that the results cannot be trusted. We submit that similar studies should be done in other types of both complementary and conventional medicine. Such studies would "borrow strength" from a large number of trials and provide empirical information to assist reviewers and readers in the interpretation of findings from small meta-analyses that focus on a specific intervention and disorder. Second, although important progress has been made lately, further research is needed to identify the dimensions of methodological quality that are important in different clinical contexts, different outcomes, and different types of trials. Finally, the relation between the probability of publication of a study and its methodological quality should be examined in more detail.

We emphasise that our study, and the trials we examined, exclusively addressed the narrow question of whether homoeopathic remedies have specific effects. Context effects can influence the effects of interventions, and the relationship between patient and carer might be an important pathway mediating such effects. Practitioners of homoeopathy can form powerful alliances with their patients, because patients and carers commonly share strong beliefs about the treatment's effectiveness, and other cultural beliefs, which might be both empowering and restorative. For some people, therefore, homoeopathy could be another tool that complements conventional medicine, whereas others might see it as purposeful and antiscientific deception of patients, which has no place in modern health care. Clearly, rather than doing further placebo-controlled trials of homoeopathy, future research efforts should focus on the nature of context effects and on the place of homoeopathy in health-care systems.

Our study powerfully illustrates the interplay and cumulative effect of different sources of bias. We acknowledge that to prove a negative is impossible, but we have shown that the effects seen in placebo-controlled trials of homoeopathy are compatible with the placebo hypothesis. By contrast, with identical methods, we found that the benefits of conventional medicine are unlikely to be explained by unspecified effects.

Contributors
M Egger conceived the study and wrote the first draft of the report. All the authors contributed to the final draft. A Shang, K Huwiler-Munsener, L Narley, S Deroi, and P Juni did the literature searches, identified eligible studies, and extracted data. P Juni advised on data extraction and quality assessment. D Pewsner helped with data extraction and classification of homoeopathic trials. A Shang, J A C Sterne, P Juni, and M Egger did the statistical analyses and contributed to data interpretation.

Conflict of interest statement
We declare that we have no conflict of interest.

Acknowledgments
We thank Fritz Grossenbacher for valuable help with literature searches. This study was funded by the Complementary Medicine Evaluation Program (Program Evaluation der Komplementärmedizin [PEK]) of the Swiss Federal Office for Public Health. We thank Marianne Amiet and Florian Mitscherlich from the PEK coordinating office and Felix Gurtner from the Federal Office of Public Health for their support. Peter Juni was supported by grants from the Swiss National Science Foundation (grants no. 3233-066377 and 3200-066378).

References