Placebo effect sizes in homeopathic compared to conventional drugs – a systematic review of randomised controlled trials

Tobias Nuhn1, Rainer Lüdtke2,* and Max Geraedts3

1Klinik Roderbirken, Roderbirken 1, 42799 Leichlingen, Germany
2Karl und Veronica Carstens-Stiftung, Am Deimelsberg 36, 45276 Essen, Germany
3Institut für Gesundheitssystemforschung, Private Universität Witten/Herdecke gGmbH, Alfred-Herrhausen-Straße 50, 58448 Witten, Germany

Background: It has been hypothesised that randomised, placebo-controlled clinical trials (RCTs) of classical (individualised) homeopathy often fail because placebo effects are substantially higher than in conventional medicine.

Objectives: To compare placebo effects in clinical trials on homeopathy to placebo effects on trials of conventional medicines.

Methods: We performed a systematic literature analysis on placebo-controlled double-blind RCTs on classical homeopathy. Each trial was matched to three placebo-controlled double-blind RCTs from conventional medicine (mainly pharmacological interventions) involving the same diagnosis. Matching criteria included severity of complaints, choice of outcome parameter, and treatment duration. Outcome was measured as the percentage change of symptom scores from baseline to end of treatment in the placebo group. 35 RCTs on classical homeopathy were identified. 10 were excluded because no relevant data could be extracted, or less than three matching conventional trials could be located.

Results: In 13 matched sets the placebo effect in the homeopathic trials was larger than the average placebo effect of the conventional trials, in 12 matched sets it was lower (P = 0.39). Additionally, no subgroup analysis yielded any significant difference.

Conclusions: Placebo effects in RCTs on classical homeopathy did not appear to be larger than placebo effects in conventional medicine. Homeopathy (2010) 99, 76–82.

Keywords: Homeopathy; Randomised placebo-controlled trials; Placebo effect; Systematic review

Introduction

Homeopathy, introduced 200 years ago, is a controversial medical system based on the hypothesis that a substance causing certain symptoms in a healthy person is also able to resolve similar symptoms in an ill patient (‘law of similars’, ‘like cures like’). Thus, in classical homeopathy, for each patient a unique, highly diluted medicine is chosen, which matches the individual patient’s symptom presentation and history.

The process of choosing the right, individually matching medicine needs not only a skilled homeopath but also an open-minded patient willing to present his or her physical and mental symptoms during the case-taking. Classical homeopathy therefore represents individuality and paying great attention to the subjective experience of the patient. Two well-known actuating variables appear to be central aspects of the homeopathic therapy: ‘meaning’ in terms of the patient’s identification with the therapy received and patient practitioner agreement in terms of the patient’s well respected subjective perspective. Meaning is regarded by Moerman as the foundation of placebo response. It thus has been hypothesised that homeopathy is something like...
a super-placebo, saying that homeopathy has the ability to evoke large nonspecific health effects.

Other aspects of homeopathic treatment, which may corroborate this hypothesis, include:

1. Typical patients of a homeopathic practice usually have a history of conventional treatment, are chronically ill\(^3\) probably with a great portion of despair and/or ambition. Despair seems to be an important actuating variable in the placebo response.\(^4\)

2. Typical patients of a homeopathic practice experience chronic illnesses with cyclic course of severity.\(^3\) Patients with such illnesses tend to enrol in clinical trials when their disease is at its worst.\(^1\),\(^4\),\(^5\)

3. Homeopathy is often not covered by health insurance and therefore a costly and exclusive, self-selected treatment, probably chosen out of a wide range of complementary treatments according to a patient’s belief system. This strengthens the meaning of homeopathic treatment.

4. Classical homeopathy is not restricted to administering medicines but also includes differing management of conventional therapies and specific life-style recommendations. Both must be considered as co-factors enhancing the effects of treatment on the patients’ health.

Homeopathy is becoming increasingly popular and constitutes an important factor of health care systems. For example, in the US the proportion of patients obtaining homeopathic care has quadrupled from 1991 to 1997.\(^7\) In the UK it was estimated that 2% had visited a homeopathic practitioner in the last 12 months\(^7\) and that annual expenditures for homeopathy reached £34.04 million (out-of-pocket £30.74 million, NHS £3.3 million).\(^8\) In Germany, the country in which homeopathy originated, a survey demonstrated that approximately 10% of men and 20% of women in the general population had used homeopathic medicines during the previous year.\(^9\) Similar percentages have been reported for Austria,\(^10\) Norway, and Denmark.\(^11\)

Meta-analyses of placebo-controlled randomised clinical trials (RCTs) show mixed results as to the efficacy of homeopathic medicines,\(^12\)\(^–\)\(^15\) with the most recent and controversial study\(^16\) finding no proof that homeopathic medicines have any effects beyond placebo.\(^17\) Walach however, suggested that these negative results do not necessarily mean that homeopathic treatment is a pure placebo, but rather that high placebo response rates may have obscured specific effects of homeopathic medicines.\(^18\) In a hypothetical study he contrasted two placebo-controlled RCTs, one with a conventional, one with a homeopathic verum. In both trials the overall treatment effects could be subdivided into four constituents: (a) statistical artifacts, (b) the natural course of the disease, (c) contextual (unspecific, placebo) effects associated with the treatment under study, and (d) specific effects of the verum.

Based on these arguments, Walach speculated that contextual effects in homeopathy are much larger than in conventional medicine, but specific effects smaller. In this case the two RCTs would have yielded puzzling results: a proof of efficacy of the conventional verum, no proof of efficacy of the homeopathic verum, but a clinical superiority of the homeopathic to the conventional verum. Moreover, this could explain why many homeopathic RCTs failed to prove efficacy – they were possibly statistically underpowered because the planning relied on placebo responses from conventional RCTs. Walach’s hypothesis has not been empirically studied yet. We therefore aimed at investigating whether contextual effects in classical homeopathy are higher than in conventional medicine. For this we compared the changes in the placebo groups of RCTs from classical homeopathy and matching conventional trials.

For simplicity these changes in placebo groups will be termed “placebo effects”, although many other effects may contribute to them, including regression the mean, the natural course of the disease, patient-physician interactions, Hawthorne effects, or concomitant treatments.

Methods

Eligibility and matching criteria

Randomised, placebo-controlled, double-blind clinical trials of homeopathy were eligible if they investigated classical homeopathy, i.e. the patients were given a single, individualised, potentised medicine according to the law of similars. We included all trials published as dissertation, abstract, or full publication before the end of 2005.

For each of the homeopathic trials identified, a set of three matching trials of conventional medicines were selected. Conventional trials were eligible if they were randomised, placebo-controlled, double-blind, and investigating chemically defined substances with a known (or hypothesised) mechanism of action. Trials of medications with an oral route were preferred.

The following matching criteria had been defined a priori: an identical general study design (parallel treatment groups or cross-over treatment), a similar study population with respect to mean age and mean severity of disease, a clearly comparable outcome measure (if possible identical, otherwise measuring the same parameter or construct), an identical way to allow or disallow for basic and concomitant medication, and an identical (or very similar) length of follow-up.

Homeopathy trials were excluded, if no sufficiently matching conventional trial could be located or relevant data could not be extracted. If the relevant data of a conventional trial could not be extracted this trial was dismissed and replaced by another trial (see below).

Study selection

Homeopathic trials were electronically searched in April 2006 in the following databases: MEDLINE, EMBASE, CAMBASE, and our own files, a database specialised in research on homeopathy which is regularly updated by an experienced bibliometician. In addition, the reference lists of systematic reviews and meta-analysis on the effectiveness of homeopathy were searched manually.

Trials of conventional medicines were primarily searched from existing systematic reviews in the Cochrane library. If more than one review was found the most recent was taken. If more than three conventional trials fulfilled all matching criteria, three trials were sampled at random. If less than
three matching trials could be located in systematic reviews, relevant trials were searched in MEDLINE.

**Data extraction**

All relevant data was extracted by the main investigator (TN). Besides the eligibility and the matching criteria mentioned above this included details on the main outcome parameter in the placebo group, i.e. mean (and standard deviation) before and after treatment/observation period, mean (and standard deviation) change during treatment/observation period, and rate of patients showing considerable improvement.

All sets of studies were classified according to several subgroup criteria: study population (children versus adults), condition treated (acute versus chronic, pain versus non-pain condition), length of follow-up (less than one week, up to eight weeks, up to twelve weeks, over twelve weeks), number of patients included in analysis (at least 20 in each trial versus at least one trial smaller), quality of the homeopathic trial report as assessed by the Jadad score: three or more points versus two points, (two was the minimum as only randomised double-blind trials were included).

**Outcome measures**

The primary outcome measure for each trial was the mean placebo effect, defined as the mean percentage improvement of the main outcome parameter (in the homeopathic trial) from baseline to the end of the study. If continuously scaled this was calculated by dividing the mean pre-post difference by the mean outcome at baseline. Otherwise, reported response rates were taken.

**Quality assessments**

The reporting quality of each trial was assessed by the Jadad score, which scores a study according to the presence of three key methodological features: randomisation (0–2 score points), blinding (0–2 score points) and accountability of all patients including withdrawals (0 or 1 score point). Higher scores indicate better quality. In addition we assessed subjectively how exact the matching criteria were fulfilled (good versus fair), and how well outcomes were comparable (adequate versus inadequate).

**Statistical analysis**

The primary statistical analysis was based on all matched sets qualified for inclusion. In each matched set we counted the number of conventional trials in which the placebo effect was smaller than the placebo effect of the matched homeopathic trial (which results in one of four possible outcomes: 0, 1, 2, and 3). Afterward these counts were averaged. Under the null hypothesis, that placebo effects in classical homeopathy and conventional medicine are identical, this average is expected to be 1.5, values above 1.5 indicate that homeopathic placebo effects are larger, values below 1.5 suggest that they are smaller than conventional placebo effects. Moreover, it is easy to compute the exact distribution under the null hypothesis (this follows from the fact that the count in each matched set is binomially distributed), from which we calculated exact one-sided p-values.

As a supplementary analysis we compared the actual size of the extracted placebo effects, here subtracting the mean conventional placebo effect from the homeopathic placebo effect in each matched set. Assuming normality of these differences and homogeneity of variances (both rather restrictive assumptions) we obtained one-sided p-values from a paired t-test.

Primary and supplementary analyses were not only done for the complete set of matched sets but also for each subgroup separately.

For each test the (one-sided) significance level was set at \( \alpha < 0.025 \). All statistical analysis were performed using SAS® 9.1 (Cary, NC, USA).

**Results**

In total, 179 placebo-controlled RCTs of homeopathy were identified, 35 of them on classical homeopathy. Of these 35 trials seven had to be excluded from analyses, because the relevant data could not be extracted. Another three trials were excluded as not enough matching conventional trials could be found (Figure 1). Thus 25 RCTs, each with three matching conventional trials, were included and evaluated. All homeopathic trials were parallel group studies, none used a cross-over design. For a list of included homeopathic trials see Table 1, a list of all included conventional trials is given as a supplementary file in the internet.

These 25 homeopathic trials covered 16 different health conditions ranging from chronic fatigue syndrome to wisdom tooth extraction, atopic dermatitis, or acne vulgaris. Four trials were on migraine or tension-type headaches, three trials on childhood diarrhoea and rheumatoid arthritis (Table 1). In these 25 trials a total of 833 patients were treated with placebo, 737 patients were analysed.
In 13 matched sets the placebo effect in the homeopathic trials was larger than the average placebo effect of the matched conventional trials, in 12 matched sets it was smaller (Figure 2). Considerable differences in placebo effects (defined as a 5% difference at minimum) were found in 9 matched sets in favour of the conventional trials. This hardly exceeds the number of 1.5 trials expected under the null hypothesis. Consequently, no difference in placebo effects (defined as a 5% difference at minimum) were found in 10 matched sets in favour of the homeopathic trial, and smaller effects (defined as a 5% difference at minimum) were found in 12 matched sets in favour of the average placebo effect of the matched conventional trials, in 13 matched sets the placebo effect in the homeopathic trials was larger than the average placebo effect of the matched conventional trials. This hardly exceeds the number of 1.5 trials expected under the null hypothesis. Consequently, no difference in placebo effects could be found (p = 0.39).

If the analysis had not primarily been based on the matched sets but on matched health conditions, counting the mean number of conventional trials with smaller placebo effects than the mean placebo effect of matched homeopathic trials treating a matching health condition, an essentially identical result would have been obtained.

Similarly, no subgroup analysis showed marked differences in placebo effects between homeopathic and conventional RCTs (Table 2). Placebo effects of homeopathy tended to be higher in adults and chronic diseases, but respective p-values were considerably above the 2.5%-level. The same applied to trials with pain as the main outcome measure.

**Discussion**

To the best of our knowledge this study represents the first systematic attempt to compare placebo effects of homeopathic and conventional treatment. Our study shows no evidence of a generally larger placebo effect in RCTs treating a matching health condition, an essentially identical result would have been obtained.

Similarly, no subgroup analysis showed marked differences in placebo effects between homeopathic and conventional RCTs (Table 2). Placebo effects of homeopathy tended to be higher in adults and chronic diseases, but respective p-values were considerably above the 2.5%-level. The same applied to trials with pain as the main outcome measure.

**Table 2** Subgroup analyses

<table>
<thead>
<tr>
<th>Subgroup</th>
<th>Number of matched sets</th>
<th>Effect size</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Children only*</td>
<td>6</td>
<td>0.83</td>
<td>0.21</td>
</tr>
<tr>
<td>Adults only*</td>
<td>10</td>
<td>1.90</td>
<td>0.33</td>
</tr>
<tr>
<td>Chronic diseases</td>
<td>19</td>
<td>1.84</td>
<td>0.02</td>
</tr>
<tr>
<td>Acute diseases</td>
<td>5</td>
<td>0.80</td>
<td>0.24</td>
</tr>
<tr>
<td>Trials measuring pain</td>
<td>7</td>
<td>2.00</td>
<td>0.18</td>
</tr>
<tr>
<td>Trial duration &gt;12 weeks</td>
<td>12</td>
<td>1.67</td>
<td>0.70</td>
</tr>
<tr>
<td>Patient numbers</td>
<td>12</td>
<td>1.67</td>
<td>0.70</td>
</tr>
<tr>
<td>High methodological quality</td>
<td>12</td>
<td>1.37</td>
<td>0.65</td>
</tr>
<tr>
<td>High matching quality</td>
<td>17</td>
<td>1.77</td>
<td>0.28</td>
</tr>
</tbody>
</table>

* Trials with mixed populations of adults and children excluded.
in classical homeopathy than in standardised conventional medicine. As a result, we cannot support the hypothesis that homeopathic RCTs are statistically underpowered due to overoptimistic placebo effect assumptions. This however does not necessarily mean that all RCTs of classical homeopathy are adequately powered, nor does it mean that classical homeopathy is a placebo therapy. Such conclusions would go far beyond the scope of our analysis.

Originally, Walach’s hypothesis was not formulated for homeopathy alone but for a wide range of procedures in complementary and alternative medicine. To our knowledge, our study is the first to approach this hypothesis in the context of classical homeopathy. It has however, been researched in acupuncture: Several RCTs indicate that an individualised acupuncture treatment has larger health effects than a conventional standard therapy, as has a sham acupuncture. These include large German acupuncture trials on migraine, osteoarthritis of the knee, and low back pain. Standard therapies were complex in these studies, but most components had already been shown to be effective beyond placebo. Consequently, one might conclude that sham (placebo) acupuncture has larger effects than a conventional placebo in the same disease.

Such studies are rare in homeopathy. We are aware of only two RCTs which systematically investigated context effects in homeopathy. Both did not directly compare placebo effects in homeopathy and conventional medicine. Steinsbekk’s RCT compared four groups of patients with upper respiratory tract infections (URTI). Its core was a placebo-controlled trial on three different homeopathic medicines, which were individually chosen by the patients themselves according to a pre-defined symptom list. Two further groups served as controls, a classical homeopathic treatment delivered by a trained homoeopath, and a routine care treatment without any homeopathic medication. As measured by an overall disease severity score there was no difference between the verum and the placebo homeopathic self medication in this trial, suggesting no specific effect of the homeopathic medicine. However, the classical homeopathy group performed significantly better than the conventional care group suggesting considerable contextual effects of classical homeopathic treatment.

In a similar study, Fisher randomly allocated 75 dermatitis patients to four treatment groups. Two groups formed a double-blind placebo-controlled RCT of classical homeopathy. A third group received open homeopathic treatment with the individually chosen homeopathic verum, and the last group was awaiting list control. In this trial the patient numbers were small, leaving the interpretability of the results rather limited. Its message however seems to contradict Steinsbekk’s results: the biggest improvements were observed in the standard care group, the smallest in the open verum group.

Although these differences were not statistically significant, and drop-out rates varied significantly (with the lowest drop-out rate in the waiting list control) this challenges the hypothesis that additional homeopathic treatment has beneficial contextual effects.

Several systematic reviews and meta-analyses intensively studied the effects in placebo groups of conventional medicine trials, and the extracted data may serve as an external validation to ours. Cho et al., for example found the mean placebo response in trials on chronic fatigue syndrome surprisingly low at 19.6% as an estimate which was even larger than ours (7.9%, 16.7%, and —21.6% for conventional trials and 7.3% for the homeopathic trial). In general however, the data from such reviews was not directly comparable to our data, simply because placebo effects were defined on different scales. For example, Macedo et al. found the mean placebo effect to be 21%, but this was defined as the number of patients who improved and cannot be compared to our data, where the placebo effect was defined as the percentage improvement on a continuous scale. The major conclusions from all these reviews were that the placebo effect varied considerably between different health conditions, and that trial design and type of intervention were major factors influencing the size of the placebo effect. This conclusion again validates our study, which was designed to control for these factors by matching.

Our study has several limitations. First, the rather small sample size: as only 25 homeopathic RCTs were included, the statistical power of our comparison is small and relevant effects may have been overlooked. Several other factors may have further decreased this statistical power. For example, we cannot exclude the possibility that a poor matching quality has affected our results. When we tried to match conventional to homeopathic trials we found substantial differences in study design even in the same diagnosis. This included duration and intensity of treatment, duration and severity of diseases, age and gender distributions, comorbidity, chosen outcome measures, allowances for concomitant treatments, and length of follow-up.

This heterogeneity of trials forced us to make more compromises in the matching process than planned, making the extracted placebo effects less comparable. Moreover, most of the included trials (homeopathic and conventional) were small and included only few patients in the placebo groups. Thus, the extracted placebo effects were rather crude and vague, which again affects the variability of results and the statistical power of our study.

The severest limitation derives from the calculation of the placebo effects itself. We defined placebo effects as percentage improvements from baseline, but this was a rather arbitrary definition, made to ensure that all effects were measured on an identical scale independent of the chosen outcome measure. Percentage changes however are extremely sensitive to the denominator (the baseline value in our study). In our study we took great efforts to match trials with similarly affected patients and comparable baseline values. Consequently, the extracted placebo effects should not be interpreted across but only within the matched sets.

Our assessments of methodological quality differs from others. Shang, for example, rated Jacob’s first trial on childhood diarrhoea as of medium quality (they used the term ‘low quality’). However, randomisation and blinding procedures in this trial were identical to those in
the subsequent trials,24,25 of the same author on the same topic, which persuaded us to rate this trial as of high methodological quality. Similar deviations in assessing trial methodology are not uncommon for homeopathic trials, as has been shown by Rutten,26 who compared respective assessments for several meta-analyses.

Walách’s hypothesis is not the only theory attempting to explain why placebo-controlled RCTs on classical homeopathy might have failed. Others argue that homeopaths often need several attempts until they find the best matching medicine. Hence, the problem with homeopathic trials would be the wrong choice of an individualised medicine that does not provoke the desired reaction. Our study however does not address this point, and therefore it should simply be interpreted as a test of Walách’s hypothesis rather than an exhaustive analysis of why placebo-controlled homeopathic RCTs might have failed.

There is a great need for further research regarding the efficacy of classical homeopathy, particularly research employing trial designs that allow for differentiating treatment effects into specific and unspecific (context) effects. Theories have been developed that these aren’t additive, but rather entangled similarly to what is described by the entanglement theory, which is practically used in quantum physics.55,56 Current literature emphasises that the phenomenon of specific and unspecific effects is far from fully understood and needs to be described by a more differentiating vocabulary.1,20 Although the placebo effect doesn’t appear to differ between trials of homeopathy and conventional medicine in a consistent and therefore significant degree, the placebo effect as we know it is everything but a gauged measure. Its indication to judge the effectiveness of treatments may be limited.

Conflict of interest

Rainer Lüdtke is employed by the Karl and Veronica Carstens-Foundation, a non-profit organisation dedicated to research in homeopathy and natural medicine. Tobias Nuhn and Max Geraedts declare no potential conflicts of interest.

Acknowledgements

Tobias Nuhn was supported by the Karl and Veronica Carstens-Foundation, Essen, Germany.

References


