Homeopathy: Meta-Analyses of Pooled Clinical Data

Robert G. Hahn

Keywords
Homeopathy · Study quality · Meta-analysis · Randomized controlled trials

Summary
In the first decade of the evidence-based era, which began in the mid-1990s, meta-analyses were used to scrutinize homeopathy for evidence of beneficial effects in medical conditions. In this review, meta-analyses including pooled data from placebo-controlled clinical trials of homeopathy and the aftermath in the form of debate articles were analyzed. In 1997 Klaus Linde and co-workers identified 89 clinical trials that showed an overall odds ratio of 2.45 in favor of homeopathy over placebo. There was a trend toward smaller benefit from studies of the highest quality, but the 10 trials with the highest Jadad score still showed homeopathy had a statistically significant effect. These results challenged academics to perform alternative analyses that, to demonstrate the lack of effect, relied on extensive exclusion of studies, often to the degree that conclusions were based on only 5–10% of the material, or on virtual data. The ultimate argument against homeopathy is the ‘funnel plot’ published by Aijing Shang’s research group in 2005. However, the funnel plot is flawed when applied to a mixture of diseases, because studies with expected strong treatments effects are, for ethical reasons, powered lower than studies with expected weak or unclear treatment effects. To conclude that homeopathy lacks clinical effect, more than 90% of the available clinical trials had to be disregarded. Alternatively, flawed statistical methods had to be applied. Future meta-analyses should focus on the use of homeopathy in specific diseases or groups of diseases instead of pooling data from all clinical trials.
Introduction

Homeopathy has a long tradition in European medicine but remains controversial due the unknown mechanism of action. Although the lack of such knowledge is not unique for treatments used in the clinic, the skepticism expressed by academic scientists in this case is fueled by the difficulty perceiving a logically reasonable explanation for why homeopathy would be effective [1].

A new era in the dispute between believers and non-believers began in the mid-1990s when evidence-based medicine was first popularized. A mechanism of action no longer needed to be proved as long as it could be demonstrated statistically that the therapy was effective. What ‘worked’ could be shown with large randomized double-blind trials but, more commonly, by systematic reviews and meta-analyses.

Methodology

Evidence-based medicine initiated a decade of struggle between believers and non-believers in which meta-analyses were used as the tool of analysis. All of them were based on virtually the same material, but authors arrived at different conclusions. The aim of the present re-appraisal of this period was to scrutinize the arguments used and to illustrate what non-believers rely on to advocate abandoning homeopathy in the evidence-based era.

Results

Linde and Colleagues’ Meta-Analysis

In 1997 Klaus Linde and co-workers in Munich received much attention after publishing a meta-analysis of homeopathy clinical trials in The Lancet [2]. The researchers had searched for homeopathy studies in a wide selection of databases. Out of 186 trials, the group identified 119 that were randomized placebo-controlled studies of clinical conditions. Of these studies, 89 provided data that were adequate for a meta-analysis.

When all data were pooled, the odds ratio and 95% confidence interval (CI) were 2.45 (2.05–2.93) in favor of homeopathy. After correction for publication bias, the odds ratio decreased to 1.78 (1.03–3.10). When only the 26 studies of highest quality were included, the benefit was somewhat weaker but still statistically significant, 1.66 (1.33–2.08).

The study by Linde and co-workers [2] demonstrated, with a likelihood of more than 95% CI, that homeopathy is overall a more effective remedy than placebo. The path was then opened for the unusual approach of pooling all published studies of one type of treatment regardless of disease and specific remedy used for the cure.

Study Quality

The earliest criticism of Linde et al.’s meta-analysis [2] focused on the fact that high-quality studies seemed to show weaker effects than studies of lower quality. In 1999, Linde’s group re-assessed this issue [3]. They divided the studies into subgroups instead of using a weighting system to consider differences in quality. Naturally, the use of subgroups reduced the capacity of the available clinical trials to demonstrate differences between treatment and placebo.

The 89 clinical trials were grouped according to the Jadad score, which describes the quality of clinical trials on a scale ranging from 0 to 5. Linde et al. [3] found that the strength of a meta-analysis became gradually poorer when dealing with studies of higher quality, but the relationship was not linear; the 10 studies with the highest quality score (Jadad 5) had greater strength in favor of homeopathy than those with Jadad 3 (19 studies) and Jadad 4 (11 studies). For all 6 Jadad score levels, homeopathy was still statistically superior compared to placebo.

Linde et al. [3] also divided the studies into 12 subgroups according to the group’s own Internal Validity Scale, in which the capacity to disclose statistically significant differences at each step was further reduced. Here, homeopathy was statistically superior to placebo at all quality levels except the highest, where 5 clinical trials yielded an odds ratio of 1.55 (0.77–3.10). However, the best estimate of the odds ratio did not differ from previous evaluations, and the lack of statistical significance is explained by the fact that the calculation is based on only a few studies.

In a subanalysis of 32 trials of individualized homeopathic treatment, Linde and Melchart [4] found an overall odds ratio of 1.62 (1.17–2.23) in favor of homeopathy. Based on the study with the best quality, the odds ratio was 1.12 (0.87–1.44), which is not statistically significant.

Edzard Ernst

Since 1997, attempts to invalidate Linde et al.’s [2–4] results have followed the path of excluding most of the clinical trials or, for various reasons, focused on smaller subgroups of studies. The first claim was made by the German-British physician Edzard Ernst, a former professor of complementary medicine in UK. In 1998, he selected 5 studies using highly diluted remedies from the original 89 and concluded that homeopathy has no effect [5].

In 2000, Ernst and Pittler [6] sought to invalidate the statistically significant superiority of homeopathy over placebo in the 10 studies with the highest Jadad score. The odds ratio, as presented by Linde et al. in 1999 [3], was 2.00 (1.37–2.91). The new argument was that the Jadad score and odds ratio in favor of homeopathy seemed to follow a straight line (in fact, it is asymptotic at both ends). Hence, Ernst and Pittler [6] claimed that the highest Jadad scores should theoretically show zero effect. This reasoning argued that the assumed data are more correct than the real data.
Two years later, Ernst [7] summarized the systematic reviews of homeopathy published in the wake of Linde’s first meta-analysis [2]. To support the view that homeopathy lacks effect, Ernst cited his own publications from 1998 and 2000 [5, 6]. He also presented Linde’s 2 follow-up reports [3, 4] as being further evidence that homeopathy equals placebo. Moreover, Ernst cited a book chapter [8] that will be commented upon later.

**Cucherat and Colleagues**

Another meta-analysis of pooled clinical data on homeopathy was authored by Cucherat et al. [9]. The group identified 118 randomized controlled clinical trials as being potentially evaluable, but excluded all except 17 (hence, 86% were disregarded). The most common reason for exclusion was that the primary end point was judged to be unclear. Prevention trials and those evaluating only biological effects were also excluded.

The patient outcomes were not pooled which is an uncommon approach in meta-analyses. Instead, the significance values (p) from the different studies were combined to arrive at an overall grand p. Out of 7 ways to combine such significance values the authors chose the one that was least prone to show a favorable outcome for homeopathy. The overall treatment effect in the 17 studies was still highly statistically significant in favor of homeopathy, p < 0.000036 (risk is less than 3.6 out of 100,000 that the difference can be explained by chance).

The clarity of this result was diluted by the subsequent removal of studies according to quality. Not much strength was lost when studies were removed that were open instead of blinded. When the 9 studies included that were randomized and double-blind ed but had lost less than 10% of patients on follow-up, the p value was still 0.0084 (risk is less than 8.4 out of 1,000 that the effect is due to chance). Finally, when only the significance values for the 5 studies with less than 5% loss during follow-up were pooled, the difference between homeopathy and placebo was only close enough to be statistically significant (p = 0.082, risk of 8.2 in 100 that the superiority of homeopathy over placebo is explained by chance). Cucherat et al. [9] remained skeptical about homeopathic treatment, although their data, even after most of the statistical power was removed by excluding 86% of the clinical trials, showed that the therapy is superior to placebo. Their impression was that the studies were of poor quality, a view not shared by others [2, 10]. Cucherat et al. [9] provided odds ratios along with the exclusion exercise, which is honest. However, the reasons for exclusion and the subsequent loss of analyzing power are unbalanced. For example, there is little reason to exclude half of the material (from 9 to 5 studies) just because the dropout incidence is reduced from <10 to <5%. In fact, a dropout incidence much higher than 10% would normally be acceptable in a clinical trial and can be handled by statistical methods.

**Shang and Colleagues**

The meta-analysis published by Shang et al. in 2005 [10] identified essentially the same set of clinical trials as Linde et al., although some recently published material was added. The purpose was to compare homeopathic remedies with conventional medical therapy, although the aftermath focused entirely on the clinical efficacy of homeopathy. The group identified 165 publications and excluded 60 for various reasons, one being that an appropriate match with a conventional medical treatment study could not be found. The authors also excluded cross-over studies. The final material consisted of 110 homeopathic trials and 110 using conventional medications.

No odds ratio was presented for the effect of homeopathy versus placebo in these 110 studies, although the authors mentioned that it was in favor of homeopathy. Instead, all except 21 studies were excluded, based on quality measures. Again, no statistics were provided. The authors then created a second set of exclusions, down to 8 studies, without clearly explaining why. Their final claim, after having disregarded 95% of the available clinical trials, was that the inverse odds ratio for homeopathy was 0.88 (0.65–1.19), which is not statistically significant. This means that the best estimate of the treatment effect is 1/0.88, i.e., homeopathy is 13% more effective than placebo.

Shang et al. [10] used the ‘funnel plot’ in the same way as the senior author (M. Egger) applied to Linde’s work in a book chapter 4 years earlier [8]. This is a scatter plot of the odds ratios versus the standard errors for a group of studies. Small studies are more likely to be published when they show a positive result, while such publication bias is more unlikely to occur when a study sample is large. As larger studies usually have smaller standard errors, the overall ‘true’ odds ratio is the one indicated when the regression line in the funnel plot approaches a standard error of zero. This means that the positive results of smaller studies are disregarded as they are assumed to be balanced by negative outcomes in studies that never came to press. By relying on a funnel plot for interpretation, conclusions are based on the existence of data we believe exist, although we do not know for sure.

The funnel plot is a hopeless research tool when making conclusions about treatment effects in a mixed set of medical conditions, because studies are powered according to expected treatment effects. For example, new studies of homeopathy in allergic and rheumatic diseases cannot, for ethical reasons, include a large sample due to the positive effects reported in previous publications [7, 11, 12]. In contrast, studies of conditions where the treatment effect is expected to be low or uncertain have to have a larger sample to reach adequate ‘power.’ Therefore, when applied to a mixture of diseases, it is impossible to alter the shape of the funnel plot shown by Shang et al., regardless of how effective homeopathy might be in allergic and rheumatic diseases. Treatment effects will always be poorer in the largest studies if we power them according to previous works in the field.

These calculations made by Shang et al. [10] have been widely used by academics and skeptics as well as by the Editor of The Lancet to claim that homeopathy lacks clinical effect. A critical discussion about this conclusion followed in a later issue of the journal [13–15].
Discussion

Meta-Analysis as a Research Tool

Many researchers are skeptical to the placebo-controlled randomized clinical trial (RCT) as the optimal tool to evaluate methods in complementary medicine. The RCT provides highly valid information about the efficacy of clearly defined treatments but is poorly suited to evaluate the efficiency of more complex interventions [16]. However, one cannot disregard the fact that the placebo-controlled RCT and the subsequent pooling of data in the form of meta-analyses are highly ranked scientific methods in school medicine. Their results will continue to have a strong impact on society’s opinion about the usefulness of complementary medicine. However, meta-analyses can arrive at different conclusions despite being based on virtually the same material. They are not performed according to strict methodology and are, to a variable extent, guided by creativity, interpretation, and personal bias. This is why everyone can find arguments for and against homeopathy in the meta-analyses of the pooled clinical data. The heterogeneity encourages critical reading including personal reflections about why the various authors have chosen to present their analysis in the way they do.

Extensive Exclusion of Data

Our considerations should include the fact that some studies rely on extensively excluding data. There must always be a sound balance between the scientific gain made by excluding studies and the limitations imposed by the associated loss of statistical power. Some of the works reviewed here, and in particular works by authors who are negative about homeopathy, reach their conclusions after having excluded 90–95% of the available trials. This is done with reference to quite small differences in quality, such as whether the dropout frequency is <10 or <5%, or with no reference at all. Little attention is given to the fact that the statistics then become based on much smaller groups of patients, which rapidly hampers the possibility of disclosing true differences between homeopathic and placebo treatments. Extensive exclusion exercises are normally excused by academic rigor but also constitute a tempting way for the non-believer to ruin any evidence there might be. The challenge for the researcher is to evaluate the available data and not to exclude virtually all of them. Studies of very poor quality and those that do not contain necessary data must always be excluded, but the remainder should be allowed to contribute to the conclusion, possibly after having been given graded importance depending on how well the studies have been conducted.

Another drawback of excluding a large number of studies is that the composition of the finally analyzed mix of conditions becomes very important to the conclusion. Here, one must remember that the overall conclusion made in these meta-analyses relates to the overall efficacy of a heterogeneous group of treatments for a heterogeneous group of diseases. One example of this problem is that the nonsignificant odds ratio for the effectiveness of homeopathy versus placebo presented by Shang et al. [10] seems to be due to a single study of muscle soreness in 400 long-distance runners [17]. Without this study, the result would have shown the statistically significant superiority of homeopathy over placebo. Moreover, Shang et al. [10] excluded all except 21 studies based on quality, and providing the results from all of them would have demonstrated the statistically significant benefit of homeopathy over placebo [13]. Why the number was further reduced from 21 to 8 was not explained beyond that the latter were ‘large’. Critical readers suspect that the authors played around with the study selection until eventually they found the desired result. Strong conclusions made about the usefulness of homeopathy made in a previous report [7] and in public [14] could fuel such behavior.

In the following debate Shang’s study was criticized for lack of transparency and the highly selected nature of the finally evaluated studies [13–15, 17].

Evidence versus Recommendation

Therapies should be evaluated in 2 steps. The first one is objective and summarizes the evidence for the efficacy of the therapy. The second step is to make recommendations for use. The clinical value of a treatment is then judged more subjectively with reference not only to evidence but also to scientific, ethical, economic, and practical perspectives.

Authorities often ask different persons to do these parts when formulating clinical recommendations or guidelines. The reason is that an objective evaluation of existing evidence should be carried out regardless of what the consequences might be. If the same individual performs both evaluations it is tempting to distort the evidence when other considerations disagree with the evidence. This author believes that overlap between these 2 roles has been common in the meta-analyses of the pooled clinical data about homeopathy. That is also why the conclusions are vastly different and the debate about the evidence contains a certain degree of emotion.

Distortion of the evidence is also common in society. The present review is based on a series of blog articles published in 2011 as a reaction to a summer campaign against homeopathy organized by skeptics in Sweden. One claim was that homeopathy is poorly studied. This is not true as the number of RCTs in this area is quite large. Many therapies used in clinical medicine are based on much less data. Another widespread argument, which was even adopted by politicians, is that not a single study on homeopathy shows a positive treatment effect. In reality, the majority of RCTs on homeopathy shows positive effects. A third claim was that the studies are of low quality, which does not receive support from researchers who have specifically evaluated this issue [2, 10]. A fourth point spread to newspapers by a professional academic was that the analyses by Ernst [7] and Shang et al. [10] demonstrate beyond a doubt that homeopathy is fraud and humbug. As we have seen, these publications represent a biased selection of the literature.
Ideology Plays a Part

The reader of this literature must be aware that ideology plays a part in these meta-analyses. For example, Ernst [7] makes conclusions based on assumed data [6] when the true data are at hand [3]. Ernst [7] invalidates a study by Jonas et al. [18] that shows an odds ratio of 2.19 (1.55–3.11) in favor of homeopathy for rheumatic conditions, using the notion that there are not sufficient data for the treatment of any specific condition [6]. However, his review deals with the overall efficacy of homeopathy and not with specific conditions. Ernst [7] still adds this statistically significant result in favor of homeopathy over placebo to his list of arguments of why homeopathy does not work. Such argumentation must be reviewed carefully before being accepted by the reader.

The most believable of the meta-analyses is still Linde et al.’s work from 1997 [2] along with the associated consideration of study quality [3] as the authors appear to maintain a reasonable balance between exclusion and statistical power. The follow-up analyses by Cucherat et al. [9] and Shang et al. [10] rest on such extensive exclusion of data that the conclusions are based on only a tiny fraction of the published studies. These meta-analyses are good examples of how the same data can yield results that are statistically in favor and not in favor of homeopathy, and having a negative result is most likely when making conclusions based on as little material as possible. Applying funnel plots to a heterogeneous mix of remedies and diseases is another example of playing around with data. If this approach is statistically correct, all further clinical trials would have to include the same number of patients regardless of the expected clinical effect. Alternatively, all treatments must exert the same effect. If not, the funnel plot is flawed.

The Way Forward

Meta-analyses of pooled data from the treatment of many conditions are difficult to interpret. From a clinical point of view, such analyses are warranted only if the same effect is measured [14]. They also provide many opportunities for those who strive to reach a predetermined conclusion to choose selectively. Further meta-analyses should focus on areas where homeopathy tends to be effective instead of diluting the results with data from diseases in which an effect is unlikely. Socioscientific issues are also relevant to discuss [19]. In the meantime, the issue continues to be a struggle between believers and non-believers, guided by plausibility bias [20].

Homeopathy is not a very strong remedy but seems to show better effects than placebo particularly in conditions that either are known or can be assumed to arise from the immunologic system. A guide was given in Linde et al.’s study [2] in which good clinical effects were to be found in allergic rhinitis, rheumatology, dermatology (except warts), and certain neuro-logical disorders, such as seasickness and migraine. Benefit was more uncertain or absent in asthma, surgery, gastrointestinal disorders, anesthesiology, and gynecology [2].

As already stated, it can be argued that evidence-based approaches including the RCT are not the best forms of evaluating the efficacy of homeopathic remedies, as well as the efficiency of more complex interventions in complementary medicine. Traditional therapy might be a better comparator than placebo due to the fact that complementary therapies often show a large nonspecific effect (“efficacy paradox”) [13, 16]. Recently, Mathie et al. [21] collected all current 263 high-standard RCTs of homeopathy in humans to enable future systematic reviews based on specific traits, such as the type of condition and whether placebo or other treatments serve as controls. Until alternative methods of evaluation have gained widespread acceptance, homeopaths have a lot to gain by aligning their therapeutic ambitions to the areas and treatments in which placebo-controlled studies have shown a better outcome than placebo. In reality, homeopathic treatments are usually individualized while most of the performed clinical trials are nonindividualized [13, 22]. Therefore, the practice is not evidence-based, regardless of the results of all the trials.

Conclusion

Clinical trials of homeopathic remedies show that they are most often superior to placebo. Researchers claiming the opposite rely on extensive invalidation of studies, adoption of virtual data, or on inappropriate statistical methods. Further work with meta-analyses should abandon the concept of summarizing all available clinical trials and focus on the effects of homeopathy versus placebo or other treatments in specific diseases or groups of diseases. One way to reduce future emotional-driven distortion of evidence by investigators and skeptics would be to separate the evidence-seeking process from the formulation of clinical guidelines more clearly.

Note

The author has never practiced, received, or studied homeopathy, but has worked in clinical medicine and performed traditional medical research in anesthesiology and surgery for the past 30 years.

Disclosure Statement

There was no financial support, thus there is no conflict of interest concerning this paper.
The practice is not usually individualized while most of the performed clinical trials in homeopathy are non-individualized [13, 22]. Therefore, the practice is not aligning their therapeutic ambitions to the areas and treating the condition and whether placebo or other treatments serve as controls. Until alternative methods of evaluation have gained widespread acceptance, homeopaths have a lot to gain by systematic reviews based on specific traits, such as the type of interventions including the RCT are not the best forms of evaluation of homeopathy versus placebo in perennial allergic rhinitis with overview of four trial series. Br Med J 2000;321:471–476.

As already stated, it can be argued that evidence-based approaches including the RCT are not the best forms of evaluation of clinical trials of homeopathy. J Clin Epidemiol 1999;52:631–636.


