Psychopharm

Journal of Psychopharmacology
22(5) (2008) 543–552

ISSN 0269-8811
SAGE Publications Ltd,
Los Angeles, London,
New Delhi and Singapore
10.1177/0269881108091259

Homeopathic pathogenetic trials produce more specific than non-specific symptoms: results from two double-blind placebo controlled trials

H Walach *School of Social Sciences and European Office of the Samueli Institute for Information Biology, University of Northampton, Northampton, UK; Academic Section on the Evaluation of Complementary Medicine, Institute for Environmental Medicine and Hospital Epidemiology, University Hospital Freiburg, Freiburg, Germany.*
H Möllinger *Health Center Socrates, Güttingen, Switzerland.*
J Sherr *Malvern, Worcester, UK.*
R Schneider *Department of Human Sciences, University Osnabrück, Osnabrück, Germany.*

Abstract

We conducted two parallel, blinded homeopathic pathogenetic trials conducted at two different sites to determine whether symptoms reported by healthy volunteers were significantly different for homeopathic remedies than for placebos. Study 1 used a two-armed design, testing ozone against placebo. Study 2 used a three-armed design, testing ozone and iridium against placebo. We found significantly more remedy-specific symptoms in provers taking ozone or iridium than in provers taking placebo in the three-armed trial and in both trials pooled for ozone and placebo. We, therefore, conclude that homeopathic remedies produce more symptoms typical for a remedy than non-typical symptoms. The results furthermore suggest a somewhat non-classical pattern because symptoms of one remedy appear to be mimicked in the other trial arm. This might be indicative of entanglement in homeopathic systems.

Key words

non-specific symptoms; pathogenetic trial; placebo

Introduction

The recent debate around homeopathy focuses on the question whether homeopathy is a placebo or not (Fisher, *et al*., 2005; Shang, *et al*., 2005; Walach, *et al*., 2005a; Walach, *et al*., 2005b). Although homeopaths hold that their materia medica, comprising more than 3000 remedies, is a collection of specific remedies that have to be applied according to the rule of matching individual symptoms of patients to symptoms proper to the remedy, critics contend that the whole system of homeopathy is nothing but a placebo. It may be possibly clinically effective, but the effects are due to the activation of non-specific effects through enhancement of hope, reduction of anxiety, good patient–doctor relationship, but not due to the pharmacological properties of the remedies themselves. The bone of contention is homeopathy's theoretical implausibility in the face of modern day physical and chemical knowledge because homoeopathic remedies are frequently applied in dilutions well beyond Avogadro's number. This *a priori* implausibility makes it difficult to just stick to empirical results and postpone theoretical debate, as most scientists operate according to an implicit Bayesian scheme of reasoning. In such a scheme, the odds are stacked high against homeopathy right from the start so that no matter how convincing empirical results may be, critics are unwilling to believe them, as long as a viable theory explaining them is missing (Vandenbroucke and de Craen, 2002). However, before one sets out to construct a theory, it is important to carefully scrutinize the empirical findings. To the long-term observer, the homeopathic database has two clearly distinct and contradictory features. On the one hand,

Corresponding author: Harald Walach, PhD, School of Social Sciences and Samueli Institute for Information Biology, University of Northampton, Boughton Green Rd, Northampton NN2 7AL, UK. Email: harald.walach@northampton.ac.uk

homeopathy seems to be helpful in clinical practice. This is documented by large observational studies and a host of case studies compiled in homeopathic journals over the two centuries of homeopathy's existence (Becker-Witt, *et al*., 2004; Güthlin, *et al*., 2004; Muscari-Tomaioli, *et al*., 2001; Schlappack, 2004; Steinsbekk and Lüdtke, 2005; Thompson, *et al*., 2004; van Wassenhoven and Ives, 2004; Witt, *et al*., 2005a; Witt, *et al*., 2005b). On the other hand, the experimental database of homeopathy is less convincing than practical clinical results would have us expect. Clinical trials have difficulties, overall, to really make a convincing point for the specificity of homeopathic remedies over and against placebos. Experimental models are still in their infancy, with a large body of data lacking stability and reproducibility in the hands of independent researchers (Vickers, 1999). This situation has led some theorists to postulate that homeopathic effects may be due to some non-local mechanism, which is as yet badly understood (Milgrom, 2002, 2005; Walach, 2003; Weingärtner, 2002).

We have been using the paradigm of homeopathic pathogenetic trials (HPT) to investigate the issue of specificity of homeopathic remedies, whereas at the same time being mindful of potential effects of non-locality. The latter is a non-trivial challenge for if homeopathic effects are due to some non-local process, standard ways of experimental testing are doomed to failure in the long run, as they presuppose a causal mechanism, which might not be operative in the first place (Lucadou, *et al*., 2007). To circumvent this problem, we seek ways to combine the framework of experimental testing with the as yet little understood boundary conditions of such non-local processes (von Stillfried and Walach, 2006). We report here on the first experiment coming out of a series of pilot attempts.

The experimental model used is a homeopathic pathogenetic trial. Such a trial follows the original scheme of homeopathic remedy testing introduced by Hahnemann, the founder of homeopathy, for researching the symptoms homeopathic substances produce (Hahnemann, 1811, 1982). In such a trial healthy volunteers take homeopathically prepared substances. The volunteers note the symptoms they experience during the trial, and the symptoms deemed specific are entered into the homeopathic materia medica and used for prescription in cases of illness when a patient presents with similar symptoms. Although in the initial phases of homeopathy the substances used in such trials were often crude and toxic, homeopaths have later on often used substances diluted beyond Avogadro's number, such as C30. This is a homeopathic potency, which has been diluted and succussed, that is, vigorously shaken, 30 times at a ratio 1:100, and hence contains 10^{-60} mol of the original substance, that is, is extremely unlikely to contain any material trace of the substance at all.

Although first attempts to translate such a homeopathic pathogenetic trial into an experimental setting were inconclusive (Walach, 1993; Walach, *et al*., 2001), a re-analysis of the data using grade of membership analysis showed a pattern of data suggestive of specific effects cancelling each other out in the cross-over design used (Walach and Kohls, 2005). We, therefore, embarked on a new set of pilot studies, which form the basis for the present report (Möllinger, *et al*., 2004; Walach, *et al*., 2004).

Our studies start from a double vantage point. In one hand, the criticism levelled against Hahnemann's original method likely to inflate symptoms because of desirability effects (Dantas, 1996), as well as against extant pathogenetic studies, regarding their methodology and validity (Dantas and Fisher, 1998; Dantas, *et al*., 2007) was taken into account. We used rigorous blinding and randomization procedures. On the other hand, we wanted to avoid the restrictions seen in many modern HPTs (Fisher and Dantas, 2001; Goodyear, *et al*., 1998; Walach, 1993) that used a predefined list of symptoms and hence reduced variability, which goes against homeopathy's unrestricted qualitative data capture method.

We wanted to imitate the original Hahnemannian method as closely as possible, while using more rigor in terms of control and blinding, yet keeping the diligent, qualitative, inquisitive approach that allows for a wide range of symptoms to be perceived. This follows modern recommendations (Riley, 1994a). Such a revival of Hahnemannian method has resulted in many new pathogenetic trials over the last two decades, producing new remedies that are in use among modern homeopaths (Riley, 1994b, 1995b; Schadde, 1995; Schuster, 1995; Sherr, 1998, 2002) (see http://www.dynamis.edu/eng/search.htm for a database of new HPTs, together with some indicators of methodology, availability, and source documents where applicable). When judging the usefulness of this information (Mortelmans, 1997), it is important to note the following: homeopathic methodology and epistemology is circular. Unusual and prominent symptoms produced by healthy volunteers are being used for prescribing (Riley, 1994a; Wehmeyer, *et al*., 1996). If the prescription heals the patient, the symptom, and by proxy the HPT, is verified (Walach and Schüppel, 1997). The proof of the symptom is not in the methodology of the HPT as such, but in the pragmatic verification in clinical application. Hence, criticisms of that methodology and doubts about the results of such HPTs, although certainly warranted in general, fails to really strike, as it neglects this circularity between HPTs and clinical application that is the epistemological basis of homeopathy. To highlight this point, we decided to conduct an HPT with modern remedies that have been studied only recently according to the standards of modern homeopathic methodology. We reasoned that the database for these remedies is new and collected to a comparatively similar standard (Riley, 1994a, 1995a; Sherr, 1994). This should allow for a test not only of HPTs in general but also of homeopathy in the making.

In these studies, we combined the standard method of qualitative data collection typical for homeopathic pathogenetic trials with rigorous experimental controls and quantitative analysis. Data of our pilot studies suggested both specific and non-specific effects. We also found significant differences for specific, remedy typical symptoms in one study. We attempted to replicate these findings in a new set of studies. Our hypothesis was that, while overall the number of symptoms should not be different between placebo and homeopathy groups, symp-

toms specific for the remedies tested should be different either in single studies or in the combined data set.

Method

We followed the methods described in detail in previous publications (Möllinger, *et al*., 2004; Walach, *et al*., 2004) and laid out in a protocol before commencement of the study. The study protocol was submitted to the Ethics Committee of Freiburg University Hospital, ethical advice was sought and granted.

We decided beforehand on a list of 20 homeopathic remedies, which were rather newly established, but considered well studied by the homeopathic practitioners (JS and HM) and incorporated in the Synthesis repertory (Schroyens, 1993), the largest homeopathic database widely used by classical homeopaths and practitioners (cf. Table 1). We conducted two studies run in parallel at two different sites (see Figure 1). Study 1 (director JS) used a two-armed design, testing one remedy against placebo. Study 2 (director HM) used a three-armed design, testing two homeopathic remedies against placebo. A homeopathic pharmacy received a randomization code, which determined the remedies to be studied. The remedies were chosen from this list and all researchers, study directors, supervisors and volunteers were unaware of the particular remedy used and to the allocation of volunteers to groups. Although one remedy was common to both studies, the three-armed study used an additional remedy in the third arm, thus allowing for pooling of both studies. The study flow and some design characteristics are depicted in Figure 1.

Table 1 List of the homeopathic remedies from which two (Ozone, Iridium) were randomly selected

Hydrogen	Granite
Chocolate	Tungsten met
Androctonus (scorpion)	Lac-Leoninum
Adamas (diamond)	Lac-Equinium
Germanium	Iridium
Haliaeetus (eagle)	Salmon oncorhynchus
Ozone	*Cygnus cygnus* (swan)
Marble	Lac-caprinum
Bambusa	*Brassica napus*
Limenitis-Bredowii (butterfly)	*Rattus rattus*

Remedies were freshly prepared by Helios Pharmacy as C30 potency. For each study, a separate randomization code of random numbers was dispatched to the pharmacy. Remedy containers and placebo containers, which were indistinguishable containing either medicated or unmedicated globules, were numbered and dispatched to proving directors. Placebo pills contained one drop of 90% unsuccussed alcohol. Thus, placebo and homeopathic medications were undistinguishable by taste or by any conventional means of detection. The code was kept at the pharmacy and accessible only to the director and one assistant. There was no communication with the pharmacy during the trial, and all stages of information release from the pharmacy to the researchers were logged.

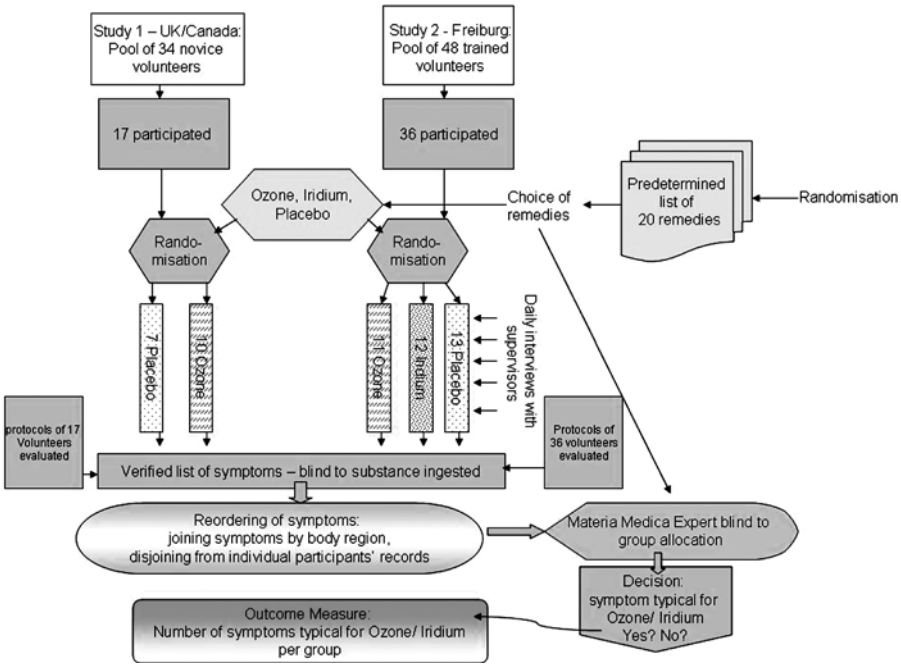

Figure 1 Participant flow and essential design features.

Participants gave informed consent. They were healthy volunteers. This was defined through a self-declaration of being healthy, not taking any conventional, recreational, or homeopathic drugs on a regular basis for any kind of medical condition, except contraception, and not suffering from any acute or chronic medical condition. There was also a quarantine on homeopathic medications of 4-12 weeks previously depending on the potency used.

Volunteers in study 1 were students of a homeopathic class learning to become homeopathic practitioners. They were familiar with the homoeopathic process of studying remedies in a pathogenetic trial, but had never participated in such a study. Volunteers in study 2 were a group of experienced homeopathic medical doctors, who had volunteered for pathogenetic trials on a regular basis before, and who knew themselves and their normal reactions very well. The volunteers did neither know the list and type of remedies nor whether they would receive placebo or a homeopathic remedy. Volunteers were instructed to report any changes in well being and any symptoms to the directors of the study. The definition of a symptom adopted was 'any unusual change from normal reaction patterns'.

As a memory aid volunteers kept an unstructured study diary in which they were to note down any change they experienced immediately, and were regularly instructed and contacted by supervisors. The supervisors contacted the volunteers on a daily basis to find out which, if any, symptoms were experienced. In a personal conversation, they helped to establish whether the reported changes were potentially worth noting or chance fluctuations of well being. This procedure followed standard criteria for HPTs, which have also been used to establish the remedy pictures of the potential proving substances (Riley, 1994a; Sherr, 1994). For instance, for each change experienced the supervisor would ask whether this change was known to the volunteer or not, if known, under which circumstances, whether the volunteer considered this unusual or not. Symptoms never experienced before were especially noted and are classically considered symptoms because of the remedy. Directors, supervisors and volunteers were blind as to the remedy chosen and group allocation at all times, until data evaluation was finished.

Volunteers were instructed to observe themselves for a baseline period of 1 week. Then they started ingesting the substances by taking five globules and dissolving them slowly in their mouth. They repeated this process at their discretion several times during the first 3 days until they experienced symptoms or changes and then stopped intake. This is necessary from the homeopathic point of view to prevent either strong symptom aggravations or else an antidoting of symptoms. The volunteers then kept closely observing themselves for another 2 weeks, reporting to the director of the respective study on a daily basis. Thus, a database of symptoms was established that was deemed likely to be due to the ingestion of the substance and different from known fluctuations in well being in a particular volunteer. The database was then reordered from head to foot according to homeopathic repertory rubrics, thus eradicating temporal and individual information. Once the database was confirmed by the study director, it was closed and sent to the study centre, which checked for plausibility and integrity.

After the database had been closed and confirmed, a third person, an expert in homeopathic materia medica and an experienced practitioner, not otherwise involved in the study was handed over the database. The study centre arranged for the name of the remedy used to be released by the pharmacy to this person only, without, however, releasing the group allocation information. Nobody else was privy to this information at this point.

The materia medica experts, a different one for each study, scrutinized every symptom of the database and determined whether it was a symptom known to be typical for the remedy tested. The computerized Synthesis repertory on Radar, Version 9.9 (Archibel, BE-Namur), was used. The result of this data evaluation procedure was the number of symptoms deemed typical for this particular remedy and symptoms untypical or non-specific symptoms experienced by each volunteer during the baseline and trial period. After this database had been established and this fact confirmed in writing to the pharmacy, the study centre received the code for the group allocation and conducted the analysis.

The question to be answered was: is the number of symptoms typical for the remedy (remedies) tested different between placebo and homeopathy groups? All statistical analyses were done using non-parametric Kruskall–Wallis tests to determine group differences, and Mann–Whitney tests to determine post-hoc differences. The main outcome was the number of symptoms typical for the remedy tested in the experimental groups during the proving phase (i.e., when the remedy was taken).

Results

Samples

Seventeen women participated in study 1. Mean age was 28.4 years (Standard deviation (SD) 8.5; range 21-58). Ten were randomly allocated to remedy, seven to placebo. Thirty-six volunteers participated in study 2. Twenty participants were women and 16 were men. Participants were on average 43.9 years old (SD 6.2; range 34-56). Eleven were randomly allocated to remedy 1, 12 to remedy 2 and 13 to placebo.

The remedies randomly chosen from the list were ozone (administered in both studies) and iridium as the second remedy given in study 2 only. The number of the symptoms typical for each remedy in all groups is shown in Table 2; examples of symptoms for each substance are given in Table 3.

The statistical analysis showed no significant differences between the numbers of typical symptoms in study 1. For study 2, the Kruskall–Wallis analysis showed significant differences for symptoms typical for iridium during baseline ($P = 0.037$), the difference being between the iridium and the

Table 2 Number of symptoms typical and untypical for each remedy in all groups during baseline and trial period (mean, standard deviation [SD], and range)

	Study 1			Study 2			
	Ozone ($n = 10$)	Placebo ($n = 7$)	Total ($n = 17$)	Ozone ($n = 11$)	Iridium ($n = 12$)	Placebo ($n = 13$)	Total ($n = 36$)
All symptoms baseline	37.2	21.4	30.7	5.1	7.2	4.6	5.6
	(37.6)	(15.5)	(30.8)	(4.3)	(5.0)	(6.0)	(5.2)
	3–122	3–48	0–419	0–10	0–15	0–21	0–68
All symptoms trial	118	64.1	95.8	22.3	14.5	12.4	16.11
	(88.4)	(43.1)	(76.4)	(16.9)	(9.9)	(12.1)	(13.5)
	11–297	23–140	0–419	3–59	2–34	1–39	0–68
Ozone symptoms baseline	25.4	14.9	21.1	0.4	1.2	1.1	0.9
	(24.5)	(10.4)	(20.1)	(0.7)	(1.2)	(2.4)	(1.6)
	1–82	1–30	1–82	0–2	0–3	0–7	0–7)
Iridium symptoms baseline	—	—	—	2.2	3.7	1.5	2.4
				(2.3)	(2.3)	(1.7)	(2.2)
				0–7	0–7	0–5	0–7
Untypical symptoms	11.8	6.6	9.6	2.5	2.3	2.1	2.3
(Placebo)	(13.7)	(7.0)	(11.4)	(2.7)	(2.5)	(3.3)	(3.3)
Baseline	2–40	0–18	0–82	0–6	0–8	0–12	0–12
Ozone symptoms trial	75.3	45.1	62.9	7.9	3.6	1.5	4.1
	(66.1)	(32.2)	(55.5)	(7.6)	(2.8)	(1.8)	(5.2)
	8–218	11–104	0–218	1–28	0–9	0–5	0–28
Iridium symptoms trial	—	—	—	9.7	6.9	8.6	8.4
				(7.8)	(4.9)	(8.5)	(8.5)
				2–24	1–14	0–31	0–31
Untypical symptoms	42.7	19.0	32.9	4.6	4.0	2.3	3.6
(Placebo)	(35.2)	(12.4)	(30.0)	(5.5)	(3.9)	(3.9)	(4.5)
Trial	3–91	6–36	0–218	0–13	0–11	0–12	0–31

placebo group ($P = 0.012$; Mann–Whitney). Also, there was a highly significant difference between the groups in symptoms typical for ozone during the experimental phase ($P = 0.0015$; Kruskall–Wallis). The difference was strongest between the group that received ozone and the placebo group ($P = 0.0008$; Mann–Whitney), but a lesser significant difference was also found between the iridium group and the placebo group ($P = 0.02$; Mann–Whitney).

When the data from both studies were pooled for the groups receiving ozone and placebo, clear significant differences emerged. These are shown in Figure 2. Mann–Whitney tests showed a significant difference between the groups for symptoms typical for ozone during the proving phase ($P = 0.011$). Other parameters, which were used for an exploratory analysis also yielded significant differences, namely the total number of symptoms between groups ($P = 0.045$) and the total number of ozone symptoms during proving ($P = 0.026$).

Discussion

We adopted a strict experimental protocol to study the effects of two newly established homeopathic remedies in 30C on healthy volunteers in two studies similar in methodology, but different in location and with different volunteers and study directors. We found clear and significant differences in symptoms typical for the remedy tested in the second study and in the combined data set. It is worthwhile noting that we combined the careful qualitative data analysis typical for homeopathic pathogenetic trials with quantitative evaluation. This methodology presupposes that the remedy in question is reasonably well known. Although we used newly established remedies, such as ozone and iridium, they are considered reasonably well known meanwhile. Qualitative analysis tries to document as many symptoms purportedly triggered by the ingestion of the remedy as possible. This yields a wide array of idiosyncratic symptoms and reactions. By matching these symptoms with the known remedy picture of that remedy from the materia medica, the main outcome variable of this study is derived: the number of symptoms typical for the remedy tested. Note that this is a kind of meta-variable, which would allow for combining different experiments conducted with different remedies.

The results we observed are very unlikely to be due to experimental errors or biases. The whole study was carried out strictly blinded for all parties involved. Neither study directors, nor volunteers, nor the personnel dealing with data handling had any knowledge of the substances tested or of the

Table 3 Examples of symptoms(*) reported by volunteers after ingestion of substances, before unblinding and without the knowledge of substances taken

	Some symptoms typical for ozone
Mind, Psyche	Immediately after ingestion, desire to breath deeply. Emotionally happy, feeling of more energy (NS 1)
Mind, Psyche	Immediately after ingestion bitter taste, with a transient feeling of aversion against the globules, afterwards very tired (NS 1)
Mind, Psyche	During the day more physical unrest than usual, had to get up more often, could not stay seated (NS 2)
Mind, Psyche	Despite a lot of work during the day very fit and balanced (OS)
Mind, Psyche	Feeling of lightness, like "flying", a bit dizzy (NS)
Mind, Psyche	Feeling strongly sedated, as if I had taken a neuroleptic medication (NS 3)
Mouth	Ingestion: dry gum, deep inhalation (NS 3)
Mouth	Bitter taste in mouth (IS)
Rectum	Hard stool, violent (OS 3)
Stool	Hard stool, with blood
Stool	Diarrhoea, sudden and violent (NS 4)
Stool	Diarrhoea, abating, but flatulence follows (NS 3)
Stool	Again diarrhoea, once (NS 2)
	Some symptoms typical for iridium
Mind, Psyche	Excitement because of several phone call, overlaying all other problems (OS)
Mind, Psyche	Mood not as continuously depressed as usual after depressing news (IS 2)
Mind, Psyche	Feeling balanced (NS)
Mind, Psyche	Mood good and balanced all day (OS)
Mind, Psyche	Mood all day excellent and very good weather (bike tour with picknick at lake) (OS)
Mind, Psyche	I say "no" to a good friend of mine's wish. But I am ruminating about this for hours (OS 3)
Mind, Psyche	Mood rather depressed, concentration not so good, dissatisfied with my work output. Trigger: problems of people close to me that occupied me a lot as well (OS 2)
Head	Headache behind forehead, eyes and below frontal part of head, dull, worse stooping (OS)
Head	On awakening, dull headache (NNH) (OS)
Ears	Feeling like pressure in ears, low humming/droning in ears, subjectively hearing became worse in the morning, better in the afternoon, worse again after 10 p.m. (NS 3)
Ears	Pressure in ears, worse right, than left; extending to mastoid (NS 2)
Larynx	Light stabbing pain in larynx, left more than right, not changed by swallowing (NS 1)
	Some symptoms experienced under placebo
Mind, Psyche	Morning, 1 h after ingestion irritated; irritation grows until afternoon (NS 2)
Mind, Psyche	Only now I realise that I have forgotten taking the globules; feel a bit stressed and forgetful; involuntarily, I stopped taking the globules; this is what happens I guess. Unusually chaotic. Feel chased around (NS4)
Mind, Psyche	In the evening sensitive against noise, irritated and angry at diner until 8 p.m., alright after things have calmed down (OS)
Vertigo	Slight vertigo on getting up and especially when showering (NS)
Head	Slight migraine, left, after quarrel with husband (OS 2)
Head	Slight pressure in head morning, after getting up (NS 1)
Larynx	Burning pain in larynx after waking up, better after swallowing (OS 2)
Throat	Enlarged lymph node left, at inception of m. sternocleidomastoideus (OS 3)
Stomach	Nausea after diner (OS 2)
Abdomen	Flatulence in abdomen after wholemeal bread for breakfast (OS 3)

NS, new symptom; OS, old symptom that is known; IS, improved symptom; indicates a healing reaction; numbers after abbreviations indicate the number of persons experiencing the symptom.

*The symptoms have been determined by individual examinations and discussions with a supervisor and by comparison with normal known status of volunteers, after being recognized as noteworthy changes from normal or expected status. By definition, all symptoms, that is, placebo symptoms as well, are unusual for the person in question. This material has been used by the materia medica expert to compare against the known symptoms of the two remedies tested and a determination was made whether the symptom is either typical for ozone or iridium or for none.

group allocation code. Randomization was completely concealed. Only the pharmacy knew about the remedies tested and the allocation code. The volunteers had no access to the code, and the directors of the study knew only the possible range of 20 different remedies from which the ones tested were chosen at random. Hence unblinding is not a reasonable explanation for our results. Even if the directors of the trial had guessed the medication, they would have had no way of biasing

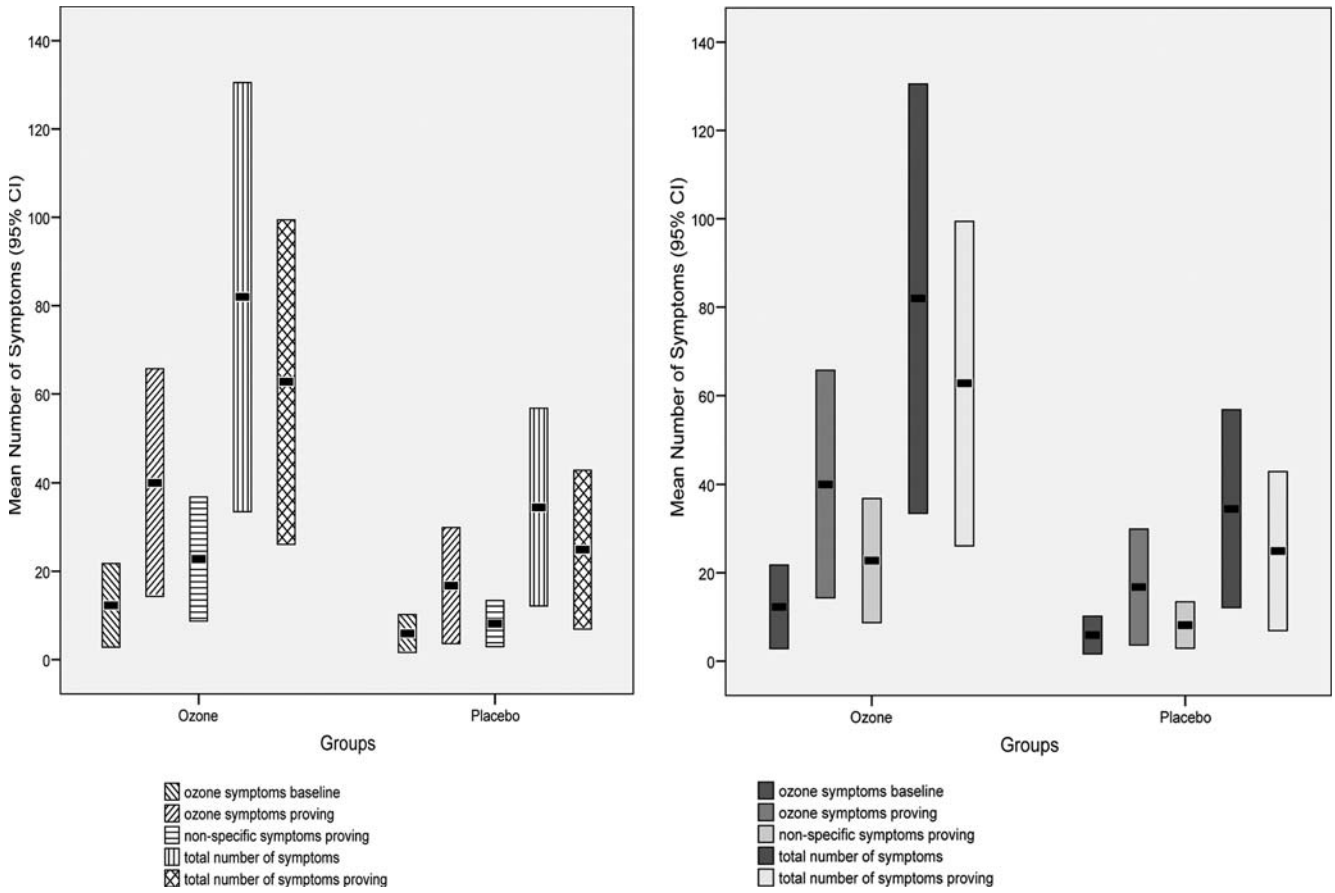

Figure 2 Combined data set: Mean number of symptoms per group and confidence intervals; significance given according to Mann-Whitney tests. Main outcome parameter: symptoms typical for ozone during proving; all other data are shown for reasons of transparency and tests are to be considered exploratory.

the data, as the allocation code was also concealed. When the name of the remedy was released to the materia medica expert, the final database of symptoms had already been deposited at the study centre; hence, no post-hoc tampering with data was possible. After data evaluation by the expert, there were no data processing steps necessary that might have unconsciously or overtly biased the results.

The fact that we have conducted two studies, slightly modified in design but using the same methodology and the same core medications with two different study directors, different volunteers, settings and locations, adds some robustness and credibility to our data and reduces the likelihood of chance findings. A clear drawback is the fact that the first study does not reach significance independently. However, it should be noted that the first study was smaller and hence had less power. More specifically, the effect between the groups for symptoms typical for ozone during proving was $d = 0.60$. This effect would have required a sample size producing twice as many symptoms to have a power of $1-\beta = 0.80$ ($\alpha = 0.05$).

Nevertheless, the effects point in the same direction and, when pooled, yield a clearly significant result. Non-specific symptoms are also experienced more frequently during the proving period by the groups receiving homeopathic remedies, not only by those receiving placebo. This might be a reflex of lacking knowledge on part of the materia medica expert, who in fact might have made mistakes, or it might have been due to the fact that in those new remedies new trials still bring out new and unknown symptoms, which then are considered non-specific because they are not yet known.

The clearest result was produced by the study that used experienced volunteers who had actually conducted many such studies before, although they produced much less symptoms overall (study 2). This underscores the fact that distinguishing symptoms from normal background noise is not a trivial task and presupposes good knowledge of one's own organism and potential reactions to situations potentially disturbing to the system. Only then is it possible to distinguish potentially new symptoms from rather trivial or known

individual reactions. Moreover, homeopathic pathogenetic symptoms are sometimes minute, hardly noticeable and often transient. This is why our methodology involved daily contacts between provers and supervisors to discuss the experiences and to decide which experiences are unusual and thus qualify as symptoms. The supervisor assumes the role of a partner listening carefully and posing questions to help decide on the experiences. The decisions are not always clear-cut as the changes observed are often very small, such as a very slight sensation of pain, or a peculiar dream, etc. (see Table 3 for examples). A volunteer has to compare the experienced symptoms with the normal state and take into consideration other events in order to determine whether a change is unusual. Also, this procedure may produce extra noise in susceptible individuals, reducing the power of an HPT to distinguish between symptoms, and one might consider the purely female sample of study 1 especially susceptible. This is visible in the comparatively high number of untypical symptoms under ozone in study 1.

Our volunteers who are from an experienced group (study 2) were clearly more successful in determining what counted as remedy symptoms. They reported far less symptoms than the novice volunteers for whom this was the first experience (study 1). The result is a clearer and significant distinction between symptoms typical for the remedy and non-specific symptoms. This difference in mean number of symptoms observed between study 1 and 2 might also be partly an artefact of recording. Symptoms can be broken down into smaller units. In a different vein, volunteers in study 2 may just not have experienced more symptoms because they were of a less volatile temperament on the study, whereas the novice volunteers were keen on experiencing a homeopathic pathogenetic trial for the first time. This line of reasoning is corroborated by the number of symptoms produced in the study with novice provers. They reported significantly more symptoms during baseline ($P < 0.001$) and during proving ($P < 0.001$) than their counterparts in study 2.

We would like to point out a noteworthy characteristic of our data: we see significantly more symptoms typical for iridium during baseline in the group later receiving iridium. This is a completely counterintuitive finding. One might be tempted to attribute it to chance, as the significance level is $P = 0.037$ and hence not convincing enough considering the amount of tests done. Nevertheless, it points to some element in our data, which one could interpret as a signature of non-locality (Walach, 2005). A similar result has been observed in another pathogenetic trial (Lewith, *et al*., 2005). Also note that during the experimental phase, although not significant, more symptoms typical for iridium were found in the group actually taking ozone than that actually taking iridium. Although with a two-armed trial such a finding can be attributed to the setting of the study and the particular style of a materia medica expert, in a three-armed trial this is not a viable explanation. For instance, a materia medica expert might be generally more prone to label a symptom as 'typical for the remedy' in the latter part of a study, thus increasing the count of those symptoms in both the control and experimental group and reducing the separability. However, it must be remembered that only a veridical decision to the best of the knowledge of the materia medica expert will be an optimal strategy under conditions of blinding and hence only an unconscious tendency is a reasonable option. Such a scoring habit can account for the result of study 1 only, but it cannot account for the results of study 2, where significantly more symptoms typical for ozone compared with placebo was found in the iridium group, but not more iridium symptoms in the ozone group compared with placebo. Such a pattern of distribution is, again, a signature for a non-local effect, we suggest. Such effects are predicted by a generalized formalism of quantum mechanics, which predicts entanglement in a generalized form under certain conditions (Atmanspacher, *et al*., 2002). It has been argued that homeopathy might be a case in question where such processes are at work (Milgrom, 2002; Walach, 2003). Our data give at least a hint that this idea is not completely off the mark.

By combining the data from two different trials where only one (blinded) remedy was common, we managed to distil at least one fact from the data: that significantly more symptoms typical for a remedy were experienced by those taking the remedy than non-specific symptoms. These typical symptoms were most prominent in the group actually taking that remedy.

One might argue that we did not verify whether our blinding was successful. This is true, but circular. Blinding is normally assessed by asking participants about assumed group assignment. Especially when side effects are prominent, these side effects are known to unblind patients in clinical trials. An HPT is designed to elicit what in a clinical context would be deemed a side effect. Hence, by the same token, it does not make sense to ask participants of an HPT about their assumption, as these assumptions would only reflect their hypotheses about assignment as inferred from the symptoms they experience, but would not allow any true conclusion as to the real success of blinding. As homeopathic remedies and homeopathic placebos are probably the only pharmacological agents that are really undistinguishable, in terms of appearance and taste, as well as in terms of any known chemical and physical properties that are measurable, any procedure to measure unblinding would be rather non-sensical. The best guarantee for a tight blinding in our study was a multi-level procedure of control, documentation, and locking up of codes that was well documented and, we believe, completely intact until the very end of the study.

We suggest that others take up our findings and improve on our method or verify our findings by their research. A few words of caution are in place. If generalized entanglement plays a role in homeopathy, then trying to pinpoint the effect as a causal one will lead to its vanishing for very clear theoretical and formal reasons laid out elsewhere (Lucadou, *et al*., 2007). Hence, detection should pursue indirect paths. Several ways have been pointed out (Walach, 2003, 2005). One obvious method was described in this article. It entails using multi-armed trials and combining data collected separately. However, as we still do not know enough about such effects, we

need further information and some more exploratory work before final conclusions can be drawn.

Some preliminary conclusions, however, can already be drawn from this study:

1) In healthy volunteers, homeopathic remedies produce more symptoms typical for a remedy than non-typical symptoms, at least sometimes.
2) Volunteers taking homeopathic remedies are more likely to experience symptoms typical for that remedy than volunteers taking placebo, at least sometimes.

Acknowledgements

The study was supported by a grant from the Samueli Institute, Alexandria, Virginia, USA. RS and HW were funded by the Samueli Institute at the time the study was conducted. The Swiss-German arm of the study (HM) was supported by the Health Center Socrates and the Socrates Foundation, Switzerland. We are grateful to all volunteers for their participation, to Nadia Bakir and Misada Vince who conducted the proving of study 1, and to Helios Pharmacy for operating the randomization and concealment procedures and for donating the remedies. We are grateful to Andrea Weninger who helped with the evaluation of symptoms.

References

Atmanspacher, H, Römer, H, Walach, H (2002) Weak quantum theory: complementarity and entanglement in physics and beyond. Found Phys *32*: 379–406.

Becker-Witt, C, Lüdtke, R, Weisshuhn, TER, Willich, SN (2004) Diagnoses and treatment in homeopathic medical practice. Forsch Komplementarmed Klass Naturheilkd *11*: 98–103.

Dantas, F (1996) How can we get more reliable information from homoepathic pathogenetic trials? A critique of provings. Br Homeopath J *85*: 230–236.

Dantas, F, Fisher, P (1998) A systematic review of homoeopathic pathogenetic trials ('provings') published in the United Kingdom from 1945 to 1995. In: Ernst, E, Hahn, EG (eds), Homoeopathy: A Critical Appraisal. London: Butterworth-Heineman, pp. 69–97.

Dantas, F, Fisher, P, Walach, H, Wieland, F, Rastogi, DP, Texeira, H *et al* (2007) A systematic review of homeopathic pathogenetic trials published from 1945 to 1995. Homeopathy *96*: 4–16.

Fisher, P, Bell, IR, Belon, P, Bolognani, F, Brands, M, Connolly, T (2005) Letter to the editor: are the clinical effects of homoeopathy placebo effects? Lancet 366: 2082.

Fisher, P, Dantas, F (2001) Homeopathic pathogenetic trials of Acidum malicum and Acidum ascorbicum. Br Homeopath J *90*: 118–125.

Goodyear, K, Lewith, G, Low, JL (1998) Randomised double-blind placebo controlled trial of homoeopathic proving for Belladonna C30. J R Soc Med *19*: 579–582.

Güthlin, C, Lange, O, Walach, H (2004) Measuring the effects of acupuncture and homoeopathy in general practice: an uncontrolled prospective documentation approach. BMC Public Health, 6.

Hahnemann, S (1811) Reine Arzneimittellehre: Erster Theil. Dresden: Arnold.

Hahnemann, S (1982) Organon of Medicine. Los Angeles: JP Tarcher.

Lewith, GT, Brien, S, Hyland, ME (2005) Presentiment or entanglement? An alternative explanation for apparent entanglement in provings. Homeopathy *94*: 92–95.

Lucadou, Wv Römer, H, Walach, H (2007) Synchronistic phenomena as entanglement correlations in generalized quantum theory. J Conscious Stud *14*: 50–74.

Milgrom, LR (2002) Patient-practitioner-remedy (PPR) entanglement: a qualitative, non-local metaphor for homeopathy based on quantum theory. Homeopathy *91*: 239–248.

Milgrom, LR (2005) The sound of two hands clapping: could homeopathy work locally and non-locally? Homeopathy *95*: 100–104.

Möllinger, H, Schneider, R, Löffel, M, Walach, H (2004) A double-blind, randomized, homeopathic pathogenetic trial with healthy persons: Comparing two high potencies. Forsch Komplementarmed Klass Naturheilkd *11*: 274–280.

Mortelmans, G (1997) What do provings prove? A critical appraisal. Homoeopath Links *10*: 201–204.

Muscari-Tomaioli, G, Allegri, F, Miali, E, Pomposelli, R, Tubia, P, Targhetta, A (2001) Observational study of quality of life in patients with headache, receiving homeopathic treatment. Br Homeopath J *90*: 189–197.

Riley, DS (1994a) Contemporary drug provings. J Am Inst Homeopath *87*: 161–165.

Riley, DS (1994b) Nicotinamide adenine dinocleotide (NAD): a proving. J Am Inst Homeopath *87*: 74–78.

Riley, DS (1995a) History of Homoeopathic Drug Provings. Philadelphia: Homoeopathic Pharmacopea Convention of the United States.

Riley, DS (1995b) Proving report – Veronica officinalis. Br Homeopath J *84*: 144–148.

Schadde, A (1995) Ozon. Eine homöopathische Studie. Munchen: Müller & Steinicke.

Schlappack, O (2004) Homeopathic treatment of radiation-induced itching in breast cancer patients. A prospective observational study. Homeopathy *93*: 210–215.

Schroyens, F (1993) Synthesis: Repertorium Homeopathicum Syntheticum. London: Homoeopathic Book Publishers.

Schuster, B (1995) Bambus. Homöopathische Prüfung und Verifizierung. Mit Kasuistiksammlung. Kronberg: Kent-Gesellschaft.

Shang, A, Huwiler-Münteler, K, Nartey, L, Jüni, P, Dörig, S, Sterne, JAC (2005) Are the clinical effects of homeopathy placebo effects? Comparative study of placebo-controlled trials of homoeopathy and allopathy. Lancet *366*: 726–732.

Sherr, J (1994) The Dynamics and Methodology of Homoeopathic Provings. West Malvern: Dynamis Books.

Sherr, J (1998) Dynamic Provings, vol. 1. West Malvern: Dynamis Books.

Sherr, J (2002) Dynamic Provings, vol. 2. Malvern: Dynamis Books.

Steinsbekk, A, Lüdtke, R (2005) Patients' assessments of the effectiveness of homeopathic care in Norway: a prospective observational multicentre outcome study. Homeopathy *94*: 10–16.

Thompson, E, Barron, S, Spence, D (2004) A preliminary audit investigating remedy reactions including adverse events in routine homeopathic practice. Homeopathy *93*: 203–209.

Vandenbroucke, JP, de Craen, AJM (2002) Alternative medicine: a 'mirror image' for scientific reasoning in conventional medicine. Ann Intern Med *135*: 507–513.

Van Wassenhoven, M, Ives, G (2004) An observational study of patients receiving homeopathic treatment. Homeopathy *93*: 3–11.

Vickers, A (1999) Independent replication of pre-clinical research in homoeopathy: a systematic review. Forsch Komplementarmed *6*: 311–320.

Von Stillfried, N, Walach, H (2006) The whole and its parts: are complementarity and non-locality intrinsic to closed systems? International Journal of Computing Anticipatory Systems *17*: 137–146.

Walach, H (1993) Does a highly diluted homoeopathic drug act as a placebo in healthy volunteers? Experimental study of Belladonna C30. J Psychosom Res *37*: 851–860.

Walach, H (2003) Entanglement model of homeopathy as an example of generalizsed entanglement predicted by Weak Quantum Theory. Forsch Komplementarmed Klass Naturheilkd *10*: 192–200.

Walach, H (2005) Entangled – and tied in knots! Practical consequences of an entanglement model for homeopathic reserach and practice. Homeopathy *94*: 96–99.

Walach, H, Jonas, W, Lewith, G (2005a) Letter to the editor: are the clinical effects of homoeopathy placebo effects? Comparative study of placebo-controlled trials of homoeopathy and allopathy. Lancet 366: 2081.

Walach, H, Jonas, WB, Ives, J, van Wijk, R, Weingärtner, O (2005b) Research on homeopathy: state of the art. J Altern Complement Med *11*: 813–829.

Walach, H, Kohls, N (2005) Grade-of-Membership (GoM) – analysis as a sensitive method for evaluating categorical data – introduction and some examples. In: Beauducel, A, Biehl, B, Bosniak, M, Conrad, W, Schönberger, G, Wagener, D (eds), Multivariate Research Strategies – Festschrift for Werner W. Wittmann. Aachen: Shaker, pp. 151–172.

Walach, H, Köster, H, Hennig, T, Haag, G (2001) The effects of homeopathic belladonna 30CH in healthy volunteers – a randomized, double-blind experiment. J Psychosom Res *50*: 155–160.

Walach, H, Schüppel, R (1997) Homöopathieforschung – Eine Taxonomie. Forsch Komplementarmed *4*: 344–347.

Walach, H, Sherr, J, Schneider, R, Shabi, R, Bond, A, Rieberer, G (2004) Homeopathic proving symptoms: result of a local, non-local, or placebo process? A blinded, placebo-controlled pilot study. Homeopathy *93*: 179–185.

Wehmeyer, A, Heger, M, Riley, DS (1996) Homöoapthische Arzneimittelprüfungen – Grundlagen und Praxis. In: Hornung, J (ed), Forschungsmethoden in der Komplementärmedizin. Über die Notwendigkeit einer methodologischen Erneuerung. Stuttgart: Schattauer, pp. 32–43.

Weingärtner, O (2002) Über die wissenschaftliche Barbeitbarkeit der Identifikation eines "arzneilichen Gehalts" von Hochpotenzen. Forsch Komplementarmed Klass Naturheilkd *9*: 229–233.

Witt, C, Keil, T, Selim, D, Roll, S, Vance, W, Wegscheider, K, Willich, SN (2005a) Outcome and costs of homoeopathic and conventional treatment strategies: a comparative cohort study in patients with chronic disorders. Complement Ther Med *13*: 79–86.

Witt, CM, Lüdtke, R, Baur, R, Willich, SN (2005b) Homeopathic medical practice: long term results of a cohort study with 3981 patients. BMC Public Health *5*: 115.