Homeopathy (2007) 96, 273–278

doi:10.1016/j.homp.2007.08.001, available online at http://www.sciencedirect.com

DEBATE

Systematic review of homeopathic pathogenetic trials: an excess of rigour?

Jeremy Sherr and Tina Quirk*

The Dynamis School for Advanced Homoeopathy, P.O. Box 652 Worcester, Worcester WR5 3WN, UK

Keywords: Homeopathic pathogenetic trial; Systematic review; Homeopathy; Human volunteers; Evidence-based medicine; Methods

We are writing in response to the study of HPTs by Dantas *et al.*[1] We congratulate the authors on their effort to study and advance provings, but we must express our concern about some of the conclusions in the discussion. The authors' proposition that "the theory linking symptoms detected in healthy volunteers to those treated in the sick is wrong," flies in the face of everything homeopathy represents.

While we believe that it is worthwhile to explore the issues raised by provings through a research model because it invites academic debate and further study, we are concerned with some of the comments in the related guest editorial. It is surprising to read the statement, "Despite a lot of effort, it remains very uncertain that HPTs yield valid results, capable of contributing to the cure of disease".[2] It seems as if the guest editor is suggesting that none of our remedies work in practice. This is perplexing because we presume that he commonly uses *Staphysagria*, *Aurum* or *Sulphur* and perhaps on occasion, some modern proved remedies. It is one thing to raise controversial questions by publishing original research, but quite another for an editorial in a respected homoeopathic journal to so strongly condemn the historical provings on which our profession is based and which have proved themselves in millions of cases.

We are keenly interested in how conventional scientific methodology can realistically study homoeopathy while maintaining its integrity and one of us (JS) has experience of conducting over 30 homoeopathic provings. From this perspective, we would like to offer the following further comments on the study and editorial.

1. It is disappointing that the authors chose to study provings from 1945 to 1995 as this is well known as the weakest period of homoeopathic provings, both in quality and quantity. Using these low quality provings as a yardstick to invalidate homoeopathic provings as a whole is seriously biased.

While admitting that HPT methodology has improved with each decade since 1945, the authors have chosen a cutoff point of 1995. In fact there was a renaissance of provings utilizing rigorous methodology following the 1994 publication of *The Dynamics and Methodology of Provings* which recommended exacting standards, including monitored pre-observation baseline, blinding of provers and supervisors, use of placebo, clear instructions to provers and supervisors, close supervision of provers, specific guidelines for selection of symptoms, and high ethical standards, as suggested by the study authors in their conclusion.[3] Unfortunately, because of the timeframe of the study, the authors have disregarded these improvements in modern proving methodology.

2. One of the most disconcerting conclusions of the authors is that historical provings, including Hahnemann's, are flawed in methodology, are unable to withstand statistical analysis for reliability and therefore produce symptoms that are unreliable in treating patients homoeopathically.

We find this a rash conclusion on the part of the authors. In the attempt to shape provings to meet RCT methodology, lies the danger of discrediting the whole body of homoeopathic knowledge that has proved successful in treating patients for over 200 years. Moreover, we have not heard of homoeopaths claiming that Hahnemann's proving symptoms of China are unreliable or that they cannot cure their Lachesis patients because the symptom data in the proving is flawed. It is notable that 70% of the reviewers in the study acknowledged that despite their poor evaluations of the methodology of the provings, they would still rely on the symptomatology of the provings to treat their patients.

*Correspondence: Tina Quirk, The Dynamis School for Advanced Homoeopathy, P.O. Box 652 Worcester, Worcester WR5 3WN, UK.
E-mail: TINQUI@aol.com
Received 13 April 2007; accepted 16 August 2007

3. The authors conclude that 'higher' quality provings produce fewer symptoms, while 'poorer' quality produce more symptoms and that fewer symptoms are indications of a better proving. As an example the authors criticize the first proving of Hydrogen (Sherr)[3] for having 5000% more pathogenic effects than a subsequent proving (Schroyens *et al*).[4]

The authors do not mention that the latter proving did not include the modern homoeopathic standards of daily supervision and provers' meeting or that the over rigorous filtering out of symptoms left only a few common symptoms, resulting in an unusable proving. There are hundreds of documented cures based on the original Hydrogen (Sherr) proving. Many symptoms from that proving have been confirmed in practice. To our knowledge there are no documented cures based on the later Hydrogen proving. Instead of rating the original proving as 5000% 'too many symptoms, the authors should rate the smaller proving as 5000% too few. In our opinion, either this proving does not conform to the 'higher' standards, or some aspects of the 'higher' standard methodology the authors advocate may be suitable for RCTs but not necessarily to HPTs.

4. The authors are 'skeptical' that 84% of the provers in the reviewed provings experienced at least one symptom and suggest that a desire to please or enthusiasm of the teachers fueled a false-positive result. They state, "If it were true one would expect many more undesirable effects of homoeopathic medicines in clinical practice".

This skepticism is not grounded in homeopathic philosophy or experience. In Paragraph 32 of The Organon, Hahnemann addresses the question of why most provers produce symptoms by pointing out the differences between natural and artificial or medicinal disease.[5] He says that in artificial disease, which includes provings, *every person* is affected at *all times* unconditionally, *regardless* of susceptibility. This was Hahnemann's experience and we have verified this repeatedly. While in some provers these symptoms may be subtle and therefore need a high degree of awareness and supervision to discern them, in our experience they prove to be most valuable clinically.

A key to understand the development of proving symptoms is to recognize that provings act as a stronger dissimilar disease (Organon para 35–42) and reflect the collective epidemic power of several people proving the same remedy at the same time. Provers do not need a high level of individual susceptibility to develop symptoms because we can control the dose while increasing our perception and awareness of symptoms. In fact, having at least one symptom from 84% of volunteers is the norm in a carefully conducted proving.

To better understand the issue of undesirable effects in clinical practice, we can look to Paragraph 156 of the Organon in which Hahnemann explains that unless a prescribed remedy is a simillimum (highly unlikely), patients will produce minor provings. This very frequent occurrence is not as apparent in a clinical setting as in supervised provings since practitioners do not monitor their patients daily after a prescription. When symptoms are noticed by the practitioner, they often are erroneously attributed to 'aggravation'.

5. The authors suggest that the "main rationale [for pathogenic trials] should presumably be human toxicity" and the guest editor echoes this by stating, "Toxins should be the first choice as candidate substances". He explains, "Choosing toxic substances is associated with advantages, including links with pharmacology and better understanding of mechanism of action. For instance symptoms detected in a recent HPT of Viscum album 30CH were probably linked with the presence of GABA in the plant. This marks a return to homoeopathy's roots: since its beginning substances toxic to humans have been selected for experimentation in healthy volunteers."

Actually, the roots of such pharmacological explanations go only as far back as Hughes. The trend suggested by the authors and editor towards low potency provings of solely toxic substances, reliance on toxicology and pharmacology and overzealous scrutiny of symptoms in order to maintain 'conventional' scientific standards at any cost, is highly reminiscent of Hughes 'Cyclopedia',[6] a book which Kent criticized as making "individualization ...quite impossible" and correctly predicted would be relegated to "the dingy corners of second hand bookstands".[7]

If Hahnemann had followed the recommendation to prove only toxic substances, we would never experience the healing benefits of *Lycopodium*, *Calcarea carbonica*, *Pulsatilla*, *Rhododendron*, *Ambra grisea* or *Lac caninum*, to mention only a few.

6. The authors state, "In fact, many important symptoms used in homoeopathic prescribing cannot be traced to HPT. Possible explanations for this include that the theory linking symptoms detected in healthy volunteers to those treated in the sick is wrong, or that the methods used in HPTs are inadequate to detect them, especially the chronic ones."

These postulations ignore another more plausible explanation. In order to produce every possible symptom from a proving, you would have to prove a remedy on a vast number of people, which is neither necessary nor desirable. One would not want to work with a proving consisting of tens of thousands of symptoms. A good proving is not about producing every possible symptom. It is about producing enough symptoms of quality so that the intelligent homoeopath can perceive a meaningful totality, which is later augmented by clinical additions.

In conclusion, it is a mistake to expect that a proving can or should attempt to be a 100% precise document. This becomes obvious when one realizes that proving the same substance on two different groups of provers will not produce identical provings. It will produce provings that are individually different but reflect the same essential meaning. This point has been well

brought out by the reproving of Cantharis and Ozone in studies by Walach, Sherr et al.[8,9]

A proving is a suggestion for a remedy and the most successful filters towards its verification are applied by examining the totality of the symptoms and by confirming and adding in clinical practice. Eliminating the majority of symptoms or characteristic single symptoms due to over scientific vigour or a concern about statistical significance or background noise, risks throwing out the baby with the bathwater. It is important to remember the proof of provings is first and foremost their clinical usability and efficiency.

Though there is much we can learn from modern research methodology, we believe that it should not be a foregone conclusion that HPT methodology should equal RCT methodology, as the authors assume. Not only are RCTs often problematic in themselves, but the very different philosophy and practice of homoeopathy require a distinctive approach. It is our experience that while modern HPTs have gained much quality by upgrading their methodology, restricting them to conventional scientific methodology can rob provings of their unique individuality.

References

1 Dantas F, et al. A systematic review of the quality of homeopathic pathogenetic trials published from 1945 to 1995. *Homeopathy* 2007; **96**: 4–16.
2 Signorini A. Guest Editorial. Finally, some light on the pillar of homeopathy. *Homeopathy* 2007; **96**: 1–2.
3 Sherr J. *The Homeopathic Proving of Hydrogen*. West Malvern: Dynamis School, 1992.
4 Schroyens F, Cecchi M, Saetonne MF, et al. Homeopathic proving of hydrogen. *Proceedings of 54th International*, Congress of LMHI, Milan, 1996.
5 Hahnemann S. *Organon of Medicine*, 6th edn. New Dehli: B Jain Publishers, 1992.
6 Hughes R, Dake JP. *A Cyclopaedia of Drug Pathogenesy*, 4 volumes. New Delhi: B Jain Publishers, 1979.
7 Kent JT. *New Remedies, Clinical Cases, Lesser Writings*. New Delhi: B Jain Publishers, 1987.
8 Walach H, Sherr J, Schneider R, Shabi R, Bond A, Rieberer G. Homeopathic proving symptoms: result of a local, nonlocal, or placebo process? A blinded, placebo-controlled pilot study. *Homeopathy* 2004; **93**: 179–185.
9 Walach H, Möllinger H, Sherr J, Schneider R. Homoeopathic pathogenetic trials produced more specific than non-specific symptoms: results from two double-blind, randomised, placebo controlled pathogenetic trials, submitted manuscript.

Authors' response: we must distinguish symptoms caused by the medicine from other symptoms

Flávio Dantas[1,2,*], Peter Fisher[3], DP Rastogi[4], Hélio Teixeira[1], José Eizayaga[5], Maria Eugenia Pulido Alvarez[6], Philippe Belon[7] and Luc Louis Maurice Weckx[2]

[1]*Universidade Federal de Uberlândia, Brazil*
[2]*Universidade Federal de São Paulo, Brazil*
[3]*Royal London Homoeopathic Hospital, London*
[4]*Central Council of Research in Homeopathy, India*
[5]*Universidad Maimonides, Argentina*
[6]*Escuela Nacional de Medicina y Homeopatia, Mexico*
[7]*Institut Boiron, France*

Quirk and Sherr raise a number of important and controversial issues concerning our systematic review of Homeopathic Pathogenetic Trials (HPTs, provings) which we will attempt to address.[1] Our 'proposition' that 'the theory linking symptoms detected through HPTs in healthy volunteers to those treated in the sick is wrong' is a hypothesis, a tentative explanation of the observed facts, and is not new. It is established that many symptoms in homeopathic materia medica derive from clinical observation, rather than HPTs. The non-systematic clinical verification of symptoms reported in HPTs is prone to several errors, such as focusing on the numerator and forgetting the denominator. The numerator may include hundreds of 'documented cures', if the medicine has also hundreds of symptoms, is frequently indicated for common symptoms/complaints and well represented in electronic homeopathic repertories. But what is the size of the denominator, how many cases were treated using the medicine without favourable results?[2] Campbell showed, in the

*Correspondence: Flavio Dantas, Department of Clinical Medicine, Federal University of Uberlandia and Federal University of Sao Paulo, Brazil.
E-mail: dantasfla@gmail.com
Received 15 August 2007; accepted 15 August 2007

early 1980s, that, for instance, if we prescribed Lycopodium solely on the basis of its provings, we would use it for a few, rather non-descript, gastro-intestinal symptoms.[3]

Quirk and Sherr assert that 1945–1995 is 'well known as the weakest period of homoeopathic provings'. We disagree: how can this be 'well-known' when ours is the first published attempt to study systematically the methodological quality of HPTs in that or any other period? The Methodological Quality Index used in our study includes randomization and blinding, traditionally used in medical studies, to reduce the effects of suggestion and selection bias. Quirk and Sherr complain that we did not extend our review by a further decade, to 2005. Ideally we would have done so, but this large international collaborative project took nearly a decade to complete, its first results were published in 1998.[4] To extend the inclusion period to 2005 would have further delayed publication.

Our systematic review showed a high incidence of pathogenetic effects reported by volunteers in HPTs which could be attributable to design flaws: more effects per volunteer were noted when the methodological quality of the trial was low. In conducting HPTs researchers need to discriminate the signal (symptoms caused by the substance being tested) from the noise (confounding factors such as the myriad events, incidents and spontaneous changes of daily life, and the symptoms and sensations related to them). In analysing HPTs the *critical task* is to separate the signal (changes caused by the medicine) from everything else.

For us the point of an HPT is to identify *true* symptoms caused by a potential medicine in healthy volunteers. This can be expressed in a simple formula: Symptoms (with medicine) − Symptoms (without medicine) = Symptoms (due to medicine).

The same logic is applied in clinical trials of new treatments, but in an inverse way (suppression of symptoms). Time is a critical factor: one cannot take an action (for instance take an homeopathic medicine), observe what happens, then go back in time and observe what would have happened if you had not taken the medicine. The closest we can come to this is to observe what happens in different people at the same time (parallel group method) or the same person at different times (crossover), or combinations of these. The use of intra-individual placebo control (crossover design) helps but does not guarantee 100% precise results. Other innovative strategies are in development.[5,6] Above all we need adequate mechanisms to cancel out 'noise' in HPTs to accept their results as valid and useful.

We do not advocate a methodological straitjacket to restrain innovation in HPTs but innovations must have a rational basis. Several biases, taken together, could contribute to the high percentage of volunteers presenting symptoms. Our results suggest that, in many cases, some or all of these symptoms would have occurred even if the volunteer had not taken the medicine. In terms of accepted knowledge, it is difficult to justify attribution of a symptom such as "got sunburn mainly on the left side of the back, although the whole of him was in the sun", reported more than 3 months after the initial dose of *Hydrogenium*, as in Sherr's report.[7] Repertories boast of the number of new symptoms they have incorporated, but never of the number of redundant ones they have eliminated, although this is necessary.[8] Quirk and Sherr may wish to suggest other criteria by which to judge what is a good HPT, but these should be justified in terms of the model they propose.

We agree with Signorini's statement that 'it remains very uncertain that HPTs yield valid results': much of homeopathic knowledge in materia medica does not come from HPTs and needs a critical review. To claim that Signorini's statement amounts to saying that 'none of our remedies work in practice' is a gross overinterpretation.

There are other important aspects to be explored in future studies, such as the sensitivity of individual volunteers (those volunteers who react strongly to a particular medicine may be the 'constitutional type'), the relationship between toxicity of the medicine and number of effects, the effects of different routes of administration, and, crucially, innovative methods to separate symptoms truly related to the medicine from those which would have occurred even if it had not been taken. We hope that all these aspects will be addressed in the future and that our systematic review will be of value to those researching these subjects.

References

1 Dantas F, *et al.* A systematic review of the quality of homeopathic pathogenetic trials published from 1945 to 1995. *Homeopathy* 2007; **96**: 4–16.

2 Stolper CF, Rutten ALB, Lugten RF, Barthels. Improving homeopathic prescribing by applying epidemiological techniques: the role of likelihood ratio. *Homeopathy* 2002; **91**: 230–238.

3 Campbell A. Lycopodium from provings. *Br Hom J* 1981; **70**: 94–99.

4 Dantas F, Fisher P. 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-Heinemann, 1998, p. 69–97.

5 Dominici G, Bellavite P, di Stanislao C, Gulia P, Pitari G. Double-blind, placebo-controlled homeopathic pathogenetic trials: symptom collection and analysis. *Homeopathy* 2006; **95**: 123–130.

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

7 Sherr J. *The Homoeopathic Proving of Hydrogen*. Malvern: The Dynamis School of Homoeopathy; 1993, no ISBN.

8 Rutten ALB, Stolper CF, Lugten RF, Barthels RWJM. Repertory and the symptom loquacity: some results from a pilot study on likelihood ratio. *Homeopathy* 2004; **93**: 190–192.