Validating Homeopathic Provings: A Pilot Study

Research Team:

Todd A. Hoover, MD, DHt
Robbert van Haselen
Joyce Frye, DO, MSCE
Ashley Ross, MD
Stephen Messer, ND
Miriam Wheeler, ND

Background

Homeopathic drug provings are experimental studies on homeopathically prepared substances given to healthy subjects with an aim to gain some understanding of the potential usefulness of the investigational substance as a homeopathic drug for human use. In other words, the homeopathic proving is the prime drug discovery tool for homeopathic medicines.

Samuel Hahnemann, who is widely recognized as the originator of homeopathic medicine, conducted proving experiments on at least 65 medicines over his lifetime[1]. He asserted that this empirical research was essential to understanding the use of a medicine. *"If one has tested a considerable number of simple medicines on healthy people in this way... then one has for the first time a true materia medica: a collection of the authentic, pure, reliable effects of simple medicinal substances in themselves; a natural pharmacopoeia*..."[2]

Many homeopathic experts consider homeopathic provings, sometimes called Homeopathic Pathogenetic Trials (HPTs), to be completely valid. This argument is based upon literally hundreds of years of verification of proving results through clinical cases. Clinicians often point to the various homeopathic repertories as sufficient evidence of the reliability of proving results in clinical application. However, no systematic evaluation of the validity of the proving method has ever been conducted.

On April 20-21, 2015, the U.S. Food and Drug Administration (FDA) conducted a formal inquiry to evaluate the regulatory framework for homeopathic medicine in the United States. During those two days, representatives from the FDA acknowledged the use of homeopathic provings as a means of drug evaluation, but pointed to randomized controlled trials as their tool of reference for drug evaluations. On September 21, 2015, the U.S. Federal Trade Commission (FTC) held a public workshop to evaluate the marketing of over-the-counter homeopathic products in the United States.[3] During this discussion, several experts testified on the lack of validation of homeopathic provings and the inability for such studies to be linked to clinical effects of medicine. So one might conclude that while many homeopaths accept the soundness of homeopathic provings without question, the FDA opinion appears to be that

the mechanism has not been fully tested, and the feedback from the FTC seemed to point to the lack of validity of such tests.

The Homeopathic Pharmacopeia Convention of the United States (HPCUS) currently uses homeopathic provings as one source of data to evaluate new homeopathic drugs for monograph purposes. The HPCUS updated its established guidelines for the methodology and analysis of modern provings in 2012. These guidelines help clarify minimum thresholds for personnel qualifications, methodology, data control, analysis, ethical conduct and legal responsibilities when conducting provings for monograph consideration. According to Guidelines on Provings for the HPCUS, sufficient outcome of a homeopathic proving is established through expert consensus that includes evaluation of proving methodology, data management, and the establishment of a homeopathic drug picture that would permit confident selection of the medicine for clinical use. While these parameters are derived from well-recognized homeopathic sources, few studies have been conducted to clarify the predictive capacity for efficacy based on these types of findings.[4,5,6,7,8]

To begin an evaluation of the validity and reliability of homeopathic provings as a drug discovery tool, the HPCUS partnered with Vithoulkas Compass to conduct a pilot study. We began the research with a simple hypothesis: Provings produce Valid Predictors (Markers or Tests) by which a clinician can select a medicine resulting in successful treatment outcomes for a patient. To satisfy this query, we would look specifically to two questions: (1) Is there any "gold" in the proving outcomes? and (2) If there is gold present, how rich is the ore?

Methods

The methods for the study were fairly simple. Take the results from well-conducted provings and compare them with the rubrics used in cases that had positive outcomes after treatment with the medicine under study. Look specifically at those cases where rubrics from the provings were used to select a medicine. Then determine how likely those rubrics would predict a positive outcome for the patient.

First, we looked for provings that had been conducted in a well-documented manner. Our hope was to locate a proving of a somewhat frequently used medicine that had been conducted within the past 15 years and, if possible, include an older proving of the same substance. After a search through National Center for Biotechnology Information (Pubmed), Encyclopedia Homeopathica, Provings.info, homeopathic provings database, and the Clificol database, we identified several candidate provings and then evaluated their methodology and outcomes according to the HPCUS standards.

We settled on *Arsenicum album* as the medicine for the study. Provings were conducted by Samuel Hahnemann on *Arsenicum album* between 1796 and 1828 with ten provers, and 1,079 symptoms were published in his Materia Medica Pura.[9] These symptoms were expanded to 2,873 with Allen's inclusion of toxicology data in his encyclopedia. Subsequently, Allen's Encyclopedia generated 1,332 (estimated) rubrics in the Synthesis TE Repertory used for this study. Möllinger, Schneider, Wallach conducted a modern proving of *Arsenicum album* that was by published in Forschende Komplementärmedizin in 2009. This proving used 15 provers of whom eight received verum and seven received placebo. Fifty

three raw symptoms were produced from the proving, and a physician from South Africa (Prof. Ashley Ross), who is experienced and competent in conducting provings, translated those raw symptoms into 272 rubrics for this study.

The CHOES, Ltd Company that developed the Vithoulkas Compass database for homeopathic cases provided access to de-identified cases. We looked at 295 cases of patients aged 2 – 90 years who were treated in the period August 31, 2014 - September 1, 2016 , and had a follow-up with outcome recorded between 18 and 180 days after the initial visit.

Results

The patients were 65% female; 35% male. They were treated in 28 different countries distributed in every continent except Africa. The clinicians for those 295 cases used an average of 17.1 rubrics per case (range 4 – 225) with a total of 2,222 unique rubrics used for treatment selection for the group. The average follow-up interval was 65.8 days (range 19 – 178 days). The most common chief complaints recorded in the cases included anxiety, depression, insomnia, allergy, asthma, headache, various pains, gastrointestinal disturbances, palpitations, and skin rashes. Mental emotional disorders accounted for 22.8% of all cases where chief complaint was recorded (58.0%). The majority (58%) of cases were treated with *Arsenicum album* in 30 – 200 C potencies. All cases used an outcome scoring of (1) No Improvement, (2) Small Improvement, (3) Moderate Improvement, and (4) Large Improvement. A Large Improvement is defined within the software used by the practitioners as _____. We chose to use only cases that had Large Improvement as truly improved with treatment.

To correlate the proving rubrics to the case rubrics used to select the treatment of *Arsenicum album*, we looked at the overlap between the provings and the cases. Over 90% of the cases used some rubrics from either the Historic proving by Hahnemann (85% of all cases) or from the Modern proving (51% of all cases). 46.4% of cases included rubrics from both provings. Seventy-two out of 272 rubrics (26.5%) in the modern proving and 289 out of 1,332 rubrics (21.7%) for the historic proving were used in the cases.

To evaluate the correlation of proving rubrics with case outcomes, we initially performed a conventional assessment of the average outcome score against the progressive percentage of proving rubrics in the cases. The cases ranged from 0 – 82% of rubrics from provings used in the case. By dividing all cases into seven roughly equal subgroups based upon increasing percentage of proving rubrics, we could assess if increasing percentage of proving rubrics in the therapy selection criteria was associated with increased likelihood of success or not. On a 0 – 3 point scale (0 = no improvement, 3 = large improvement), the average outcome for all cases is 1.73 (the seven groups ranged 1.58 – 2.05 with maximum improvement in the group with 20 – 25% of rubrics from the provings). Using these seven groups of cases, analysis shows no difference in outcomes due solely to proportion of proving rubrics in the case. According to this result alone, we would have to conclude that provings are not correlated with treatment outcomes for *Arsenicum album*.

But analysis of proportional density of proving symptoms used in cases alone relies upon averaging of data to the lowest common denominator. Homeopathic prescribing however is based upon individuality

and particulars. If we wish to evaluate the data to find any rubrics of predictive value from the provings, we can use Bayesian statistics. This approach considers the specific rubrics, rather than simply the aggregate totals of rubrics used. The assumption here is that provings may produce some valuable data that is clouded by irrelevant background noise. The rubrics themselves are a type of diagnostic "test" for the patients who would benefit from treatment with *Arsenicum album*. Considered from this vantage point, we must evaluate the cohort of those patients who were treated and had a large improvement with *Arsenicum album* to the entire database of all cases within Vithoulkas Compass during our study period. To evaluate each individual rubric as a test, we must determine which groups of cases that are True Positives, True Negatives, False Positives, and False Negatives using the rubric as our "test" and a large improvement as a positive outcome.

With all database cases organized in this manner, we can determine the Sensitivity and Specificity of our "test" for each rubric. From there, we can calculate the Likelihood Ratio for the rubric. In evidence based medicine, the Likelihood Ratio is used to assess the value of performing a diagnostic test. The Likelihood Ratio (LR) is the probability of a person who has the disease testing positive divided by the probability of a person who does not have the disease testing positive. To calculate the LR, take the ratio of all true positive cases for the rubric to all cases that improve with *Arsenicum album* (all responders) and divide that by the ratio of all false positive cases to all cases in the database that were either treated with another medicine or did not respond to *Arsenicum album*. In other words it is the ratio of the prevalence of the rubric in the Arsenicum album responders to the prevalence of the rubric in the remainder of all cases.

When the likelihood ratios are determined using the data in our study , we find that for proving rubrics found in four or more cases, ten rubrics emerge with LRs > 1. Table 1 displays the number of cases treated with *Arsenicum album* using the rubric for diagnosis (whether improved or not), the LR, and the confidence interval. The higher the LR, the more likely the rubric will be associated with an improvement if the patient is treated with *Arsenicum album*.

Table 1

Rubric	Freq.	LR	95% CI
MIND - FEAR - alone, of being	32	**4.1**	2.4 -- 7.1
MIND - FEAR - solitude, of	11	**35.9**	15.3 -- 83.8
MIND - ANXIETY – night	7	**19.8**	7.2 -- 54.4
MIND - ANXIETY - trifles, about	7	**7.9**	2.5 -- 24.5
MIND - ANGER, irascibility - trifles - agg.	4	**10.2**	1.4 – 75.7
MIND - CONCENTRATION - cannot fix attention	4	**9.1**	2.2 -- 37.2
STOMACH - THIRST - small quantities, for	16	**6.8**	2.6 -- 18.0
STOMACH - THIRST	15	**2.3**	1.12 -- 4.61
ABDOMEN - PAIN - Hypochondria - right	5	**6.2**	1.5 -- 24.9
EXTREMITIES - TREMBLING	4	**15.1**	3.6 – 62.9

Further analysis of this data will allow us to predict the probability that a given patient with that rubric might respond to treatment with *Arsenicum album*. By using dataset of all cases, we can determine the prevalence of "disease" in the whole population of patients who seek homeopathic treatment. When using only Large Improvement as a positive outcome, the prevalence or pre-test chance of patients responding to *Arsenicum album* is 0.58%. If we consider either Large or Moderate improvement to be a positive response to treatment , the prevalence of *Arsenicum album* sensitive patients with increases to 1.44%. The true prevalence likely lies between these two numbers.

By knowing the pretest probability (chance) or prevalence, we can calculate the post-test probability using the LR. The following calculations describe the process:

1. Pretest probability / (1 - pretest probability) = pretest odds.
2. Pretest odds X LR = post-test odds.
3. Post-test odds / (1+ post-test odds) = post-test probability (chance).

Clinicians unknowingly use post-test probability in the process of diagnosis day in and day out. For example, an emergency room physician may assess a patient with migrating abdominal pain. Based on the population that comes to the ER with this type of symptom, the prevalence of appendicitis may be as high as 1%. The migrating nature of the pain might increase that prevalence estimate to 3 percent. If the patient also has fever, the chance goes up again. Now if you add abdominal guarding and elevated white blood count to the picture, the chance of appendicitis might rise to 50% or so. This means that the patient has a fifty-fifty chance of appendicitis based upon those 4 diagnostic tests.

The same process occurs in homeopathic practice. Imagine you have a patient with a strong fear of being alone (LR = 4.1 for *Arsenicum album*) and clear thirst for water in sips frequently (LR = 6.8) and pain in the right hypochondria region (LR = 6.2). In this case, the sequential analysis of the chance of responding well to *Arsenicum album* for this patient goes from 0.58% before any evaluation of symptoms and rises to 50% based upon those three symptoms alone. A similar process with three other symptoms like difficulties with concentration where the patient cannot fix the attention, anxiety about trifles, and generally increased thirst raises the pre-test chance of 0.58% before evaluation to a 46.5% chance of responding favorably to *Arsenicum album* treatment.

Assessed in this way, we can see that these ten rubrics from the provings demonstrate clear predictive value for the use of *Arsenicum album* therapeutically. Any rubric with an LR > 1 from a proving is evidence that some gold is available in the proving. The analysis of the proving through actual cases can help separate the gold from the background data to provide an improved therapeutic guide for clinicians. In our study, 14.5% of all proving rubrics used in cases showed an LR > 1. The historic proving of Hahnemann showed a greater richness with 16.5% of the rubrics studied showing positive LRs, while the modern proving was associated with 12.5% of the rubrics studied.

Discussion

In conclusion, our data is consistent with our hypothesis that the provings of *Arsenicum album* contain useful clinical predictors for the successful therapeutic use of the medicine using fairly strict outcome criteria. Therefore, the homeopathic proving when carefully conducted in the fashion described by Hahnemann may be a useful drug discovery tool. However, at the present time there is no validated method to extract the “gold” rubrics from all of the rubrics generated in a homeopathic proving.

Some potential confounders in this study include the lack of inter-rater reliability testing in the outcomes measures, the likely presence of confirmation bias related to “pre-judging” or selecting rubrics based upon a desired therapeutic choice, and the high degree of variance in the number of rubrics used by various practitioners for therapeutic selection. Further research is needed to test the

outcomes of this pilot study in a prospective fashion. In addition, research into discernment of factors that can separate the rubrics that will have significant LRs for treatment outcome from less reliable rubrics in the proving could be of immense benefit when evaluating the potential utility of new medicines.

We wish to thank HPCUS and the Vithoulkas Compass team for their tremendous support in the project. Special thanks is due to Prof. Ashley Ross for his work on translating the modern proving into rubrics and to Dr. Lex Rutten for his guidance on the use of Bayesian statistics.

---

[1] Hahnemann S. *Materia Medica Pura.* Translated by R. E. Dudgeon. New York: Boericke and Taffel. 1880. Pp 769.

[2] Hahnemann S. *Organon of Medicine.* 1810. Translated by Kunzli and Naude . Traverse City:Cooper.1978. Verse 143.

[3] Homeopathic Medicine and Advertising: An FTC Workshop. September 21, 2015. Transcripts online at: https://www.ftc.gov/system/files/documents/public_events/644921/homeopathic_medicine_workshop_transcript_9-21-15.pdf. Last downloaded May 29, 2017.

[4] Bell, I.R. Evidence-Based Homeopathy : Empirical Questions and Methodological Considerations for Homeopathic Clinical Research. *J Altern Complement Med*. Spring 2003, Vol 96,n°1.

[5] Stolper CF, Rutten ALB, Lugten RFG, Barthels RJWM. Improving homeopathic prescribing by applying epidemiological techniques: the role of LR. *Homeopathy* 2002;91, 230-238.

[6] ALB Rutten, CF Stolper, RFG Lugten, RJWM Barthels. Assessing likelihood ratio of clinical symptoms: handling vagueness. *Homeopathy* 2003; 92 182-186.

[7] Van Wassenhoven M. A step to Evidence Based Homeopathy: Veratrum album. *Revue Belge D'Homoeopathie* 2004 n°2, 230-241.

[8] Van Wassenhoven M. Clinical verification in homeopathy and allergic conditions. *Homeopathy* (2013) 102, 54-58.

[9] Samuel Hahnemann. Materia Medica Pura. Vol. ii, 3rd edit., 1833 translated by R.E. Dudgeon: Hahnemann Publishing House (London) 1880.