Statistical Tools for Alternative Research in Plant Experiments

Maurizio Brizzi[1] and Lucietta Betti[2]

Abstract

In this paper we describe the set of statistical tools applied for evaluating a series of experiments concerning the effects of homeopathic treatments in simple plant models, involving wheat seeds and seedlings and tobacco plants. The homeopathic treatment used in our experiments is arsenic trioxide, diluted in a decimal scale and dynamized k times (with k varying from 5 to 45). Since the most significant results have been achieved with As 45dH (45-th decimal potency of As_2O_3), we report a brief summary of these results. The statistical analysis was performed by using parametric and non parametric tests, and Poisson distribution has an essential role when dealing with germination experiments. Finally, we describe some interesting results related to the changes in variability, which seems to be a primary target of homeopathic treatment effect.

1   Introduction

The effectiveness of homeopathic treatments has been discussed for many years among physicians and researchers, still being an open topic in the scientific community (Shang et al., 2005; Lüdtke and Rutten, 2008). This research field needs a careful statistical approach, in order to add scientific evidence to the personal opinions of "experts", either in favour or against this therapy. One of the most repeated criticisms is the presence of a placebo effect; a suitable way to avoid this kind of effect is to define experimental models where "patients" are not human beings, but plants or microorganisms (Betti *et al.*, 2003a and 2008). Furthermore, relatively simple model systems have the advantage of a more direct treatment/effect relationship, and give the opportunity of collecting large data samples for structured statistical analyses. It is useful to remark that a rigorous scientific approach is particularly important when dealing with ultra-high dilutions (beyond the Avogadro number), where no molecule of original substance is still present in final treatments (Elia *et al.*, 2007).

---

[1] Department of Statistical Sciences, University of Bologna, Italy.
[2] Department of Agri-Environmental Sciences and Technologies, University of Bologna, Italy.

In such a context, characterized by research patterns where new possible ways of investigation may be opened at any time, and little previous literature is available, the choice of suitable statistical tools has to be carefully reconsidered for every series of experiments. The presence of a professional statistician within research groups in homeopathy and other complementary medicines, although strongly recommended, has very seldom been observed up to now. We think that research work based on plant models, if it gives the appropriate weight to statistical analysis and interpretation of results, could yield important contributions to the understanding of the mechanism of action of homeopathic medicines.

2   Plant model description

During our series of experiments (1991-2009) different plant models were set up; specifically, we considered wheat germination and growth, as well as tobacco/tobacco mosaic virus (TMV) interaction.

The first and simplest model we used is based on *in vitro* wheat germination, where a sample of seeds was placed on sterilised sand in randomly distributed Petri dishes (Betti *et al.*, 1994; Brizzi and Betti, 1999; Brizzi *et al.*, 2000, 2009). Since homeopathic treatments in human medicine are to be used in unhealthy individuals, in some of our experiments we tried to reproduce a similar "disease pattern" by previously stressing wheat seeds with material doses of arsenic trioxide (As_2O_3), thus reducing germination rate. Stressed seeds were then treated with a fixed quantity of treatment: distilled water (C, control group) or arsenic trioxide diluted and dynamized at the *k*-th decimal potency, for several values of *k* ranging from 23 to 45 (treatment groups). For instance, As 45dH is used to denote the 45-th decimal potency of arsenic trioxide. The working variable is the number of non-germinated seeds per dish, out of a fixed number of 33 seeds, after 4 days of observation.

Using the same plant species, we set up a second model, concerning wheat growth, where a different biological parameter was evaluated: stem length of seedlings after 7 days of observation (Betti *et al.*, 1997; Brizzi *et al.*, 2002, 2005). In this kind of experiment, each seed was placed in a transparent cellophane envelope, inserted in a larger cardboard envelope, so that stem and roots could develop in natural light and darkness, respectively. In this model too, non-stressed and stressed seeds were both treated with a fixed quantity of distilled water (C) or arsenic decimal potencies ranging from 5 to 45 dH.

Finally, the effects of homeopathic treatments were checked on a phytopathological model, involving tobacco plants subjected to TMV inoculation as biotic stress (Betti *et al.*, 2003b). This virus induces necrotic lesions on tobacco plants carrying the TMV-resistance gene *N*. A large number of leaf disks from TMV-inoculated plants were placed for three days in randomly distributed Petri

dishes containing the same quantity of distilled water (C), W 5dH or 45dH, As 5dH or As45dH. The leaf disks were labelled in order to distinguish the "mother plant". The working variable here is the number of necrotic lesions observed in a leaf disk three days after virus inoculation. The research was structured in eight separate experiments; in the present paper, we report the results of the three experiments involving W 45dH and As 45dH.

Due to the specific biological features of above mentioned working variables, the statistical approach was necessarily differentiated. In the following paragraphs, the statistical methods applied for studying data obtained in the above described experiments will be described in detail. The most important results will also be briefly reported. In particular, we will describe in detail all the results concerning the 45dH potency, both for arsenic trioxide (As 45dH, diluted from a mother tincture and dynamized at each dilution step), and water (W 45dH, where every dilution/dynamization step was performed using only distilled water). We chose the 45dH potency since it yielded the most important results for every kind of model.
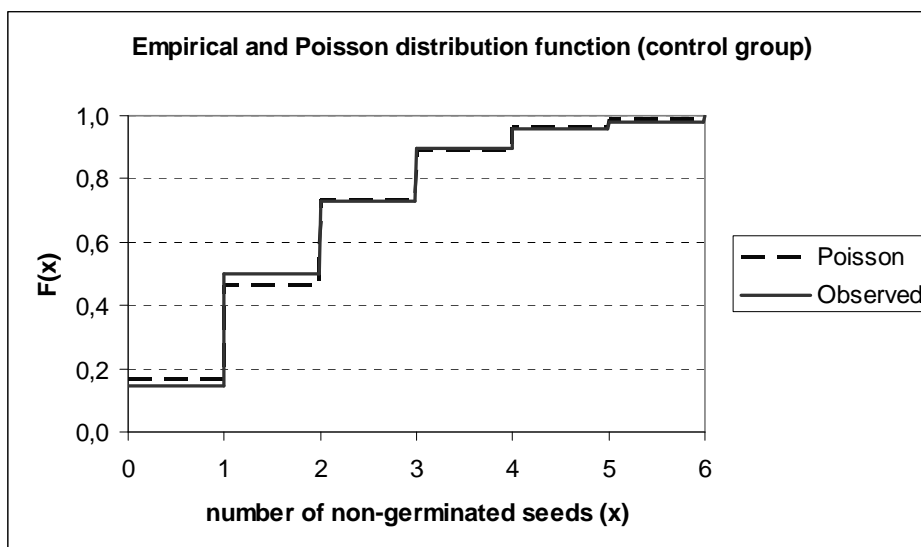
3   Poisson distribution

3.1   Theory and methods

At the beginning of our series of experiments, regarding the wheat germination model, we realized that Poisson distribution fits the number X of non-germinated seeds per Petri dish very satisfactorily. Just to give an example, we report here (Figure 1) the graphical comparison of empirical and Poisson cumulative distribution function for the control group: we can clearly note that theoretical and empirical values are very close. Nevertheless, we repeatedly checked the goodness-of-fit by means of Kolmogorov-Smirnov test (Stephens, 1974), applied to different experimental samples (both control and treatment groups), and we never detected a significant difference.

We also confirmed the adequacy of Poisson distribution for wheat germination data with a graphical method proposed by Hoaglin (1980), called the *Poissonness plot*. If we denote the values of the working variable X with x (x= 0, 1, 2, …) and the corresponding frequencies with $n(x)$, in a two-dimensional diagram we can plot x values against

$$y(x) = ln\ n(x) + ln\ (x!) \tag{3.1}$$

Figure 1: Goodness-of-fit of Poisson model to wheat germination data (Betti et al., 1994).

If the plotted points approximately follow a straight line, the Poisson distribution fits well. We checked the linearity of the plotted points by the Bravais-Pearson correlation coefficient *r(x,y)* (Brizzi and Betti, 1999). Since the *r*-values calculated are all greater than 0.83, we went on supposing that *X* follows a Poisson distribution under all the experimental situations we dealt with.

Once the goodness-of-fit of Poisson model to wheat germination data had been demonstrated, we were able to apply the specific parametric tests for this distribution. When doing pairwise comparisons, where the null hypothesis is $H_0:$ $\lambda_A = \lambda_B$, i.e. the equality of two Poisson parameters, if the samples have the same size, we can apply the following test statistic, reported by Sachs (1984):

$$\hat{z} = \frac{T_A - T_B - 1}{\sqrt{T_A + T_B}} \tag{3.2}$$

where T_A and T_B are the total number of non-germinated seeds in the two samples compared. Under the null hypothesis H_0, the test statistic (2) follows a standard normal distribution.

Since in most of our experiments control groups (C) were larger than treatment ones (T), when comparing C vs. T, we applied an exact Poisson test in the following way: if $\hat{\lambda}_C$ is the maximum likelihood estimate of Poisson parameter in the control group, and n_T is the sample size in the treatment group, we calculate the Poisson probabilities with parameter $\hat{\lambda} = \hat{\lambda}_C \cdot n_T$, which corresponds to treatment distribution under the null hypothesis $\lambda_C = \lambda_T$ Therefore, given a

significance level α, if the observed number of non-germinated seeds in treatment group, say N_T, lies in the tails of the Poisson distribution with parameter $\hat{\lambda}$, we rejected the null hypothesis.

When comparing more than two Poisson distributions in a global comparison, in a sort of "Poissonian ANOVA", we applied the test proposed by Sachs (1984). The null hypothesis is now $H_0 : \lambda_1 = \lambda_2 = ... = \lambda_k$; we used x_i to denote the total number of non-germinated seeds in the i-th sample, with t_i the total number of experiments of the same sample,

$$\hat{\lambda} = \frac{\sum_{i=1}^{k} \hat{\lambda}_i t_i}{\sum_{i=1}^{k} t_i} \qquad (3.3)$$

being the overall Poisson parameter estimate. Now, if we compute the values:

$$z_i = \begin{cases} 2(\sqrt{x_i + 1} - \sqrt{t_i \cdot \hat{\lambda}}) & if \quad \hat{\lambda}_i < \hat{\lambda} \\ 2(\sqrt{x_i} - \sqrt{t_i \cdot \hat{\lambda}}) & if \quad \hat{\lambda}_i \geq \hat{\lambda} \end{cases} \qquad (3.4)$$

the overall Poisson test statistic is the following:

$$w = \sum_{i=1}^{k} z_i^{2} \cong (H_0) \cong \chi^2_{k-1} . \qquad (3.5)$$

The test is one-tailed on the right, so we reject H_0 only for large values of w.

When dealing with multiple treatments, we began our comparison with the global test based on the statistic w; after checking that overall comparison was significant, we went on with pairwise tests.

3.2   Main results

The above described methodology, based on Poisson distribution, was applied to the *in vitro* wheat germination model (Betti et al. 1994; Brizzi et al. 2000 and 2009). First of all, we compared control samples with a global Poisson test, and the results were never significant; this result gives us important methodological information, assuring us that our model is stable and that observed differences are to be imputed to treatment effect. This was confirmed by the fact that global Poisson test becomes significant when adding treatment groups. Since we considered several decimal potencies in these works, we report here (Table 1) only the results referring to As 45dH and W 45dH with respect to control (distilled

water). It should be remembered that our working variable is the number of non-germinated seeds in a "standard trial" of 33 seeds in a Petri dish.

Looking at the table, we can immediately note that control mean values seem to be rather regular, throughout all our period of experimentation, both for non-stressed and stressed seeds. This induces us to believe that our model is quite stable, and that observed differences can be properly attributed to treatment effect. Treatment with dynamized water and arsenic does induce a significant reduction in the number of germinated seeds, showing a stimulating effect. Such an effect is considerably more significant (*p*-value constantly less than 0.01) and always reproducible when using As 45dH. Confirming our forecasts, the treatment effect was more evident when working with stressed seeds; this is the reason for using only stressed seeds in the last experiment (Brizzi *et al*., 2009).

Table 1: Wheat germination results: mean number of non-germinated seeds and Poisson test significance.

Reference	Stress		Control	As 45 dH		W 45 dH	
		n	Mean	Mean	p-value	Mean	p-value
Betti et al. (1994)	No	24	1.79	1.04	0.0021	---	---
Brizzi et al. (2000) 1st exper.	No	24	1.54	0.94	0.0056	1.44	n.s.
Brizzi et al. (2000) 2nd exper.	No	24	2.00	1.13	0.0042	1.13	0.0042
Brizzi et al. (2000) 1st exper.	Yes	24	5.03	3.25	<0.0010	3.78	<0.001
Brizzi et al. (2000) 2nd exper.	Yes	24	6.58	3.38	<0.0010	4.71	<0.001
Brizzi et al. (2009)	Yes	48	5.71	4.50	0.0044	4.90	0.0419

Legend: n = number of standard trials; Control = distilled water; As 45 dH = Dynamized arsenic trioxide; W 45 dH = dynamized water. The p-values are related to comparisons with control. In our paper published in the year 2000, we described two separate experiments.

4   Non parametric rank tests

4.1   Theory and methods

Whenever we are dealing with skewed data, which are evidently far from being normally distributed, we need to choose specific tools to make statistical inferences. When evaluating our series of experiments, we decided to apply non-parametric tests based on ranks, which means that observed data are ordered and their values are replaced by the corresponding ranks, computing the mean rank when we have identical values. When comparing the level of magnitude of two

populations, we can mainly use two distinct tests: the Mann-Whitney test for independent data and the Wilcoxon test for dependent (paired) data (see e.g. Conover, 1980). Briefly, the Mann Whitney test is based on a global ranking of sample data: if the ranks of data from one sample are considerably smaller or greater to the other sample's ranks, the null hypothesis is rejected. Indeed, the test statistics U is based on the sum of ranks of each sample; when the sample is small we do need specific tables, but if the sample is not too small (at least 20 observations in each sample), we can use a normal approximation. On the other hand, in the Wilcoxon test for paired data we need to calculate the differences of paired values (which necessarily have to be numerical), assigning a rank to such differences: the null hypothesis is rejected when there is a prevalence of differences of the same direction with respect to the other direction. This test needs specific tables which are easily available (even in the web).The Wilcoxon test for paired data is surely more powerful, but it needs a reasonable criterion for pairing data, otherwise Mann Whitney test is more adequate.

As far as our plant model system is concerned, we applied the Mann-Whitney rank sum test when dealing with wheat stem length, since the data were markedly skewed. On the other hand, when evaluating the data of the phytopathological model Tobacco/TMV, we preferred to use Wilcoxon test, after pairing leaf disks from the same plant and sharing the same rank within the plant subsamples.

Table 2: Wheat growth results: mean stem length (cm) and Mann-Whitney test significance.

Reference	Stress	n	Control	As 45 dH		W 45 dH	
			Mean	Mean	p-value	Mean	p-value
Betti et al. (1997)	Yes	150	3.17	3.94 (+24.0%)	0.001	n.d.	n.d.
Brizzi et al. (02,05)	Yes	30	6.02	7.51 (+24.7%)	0.042	7.22(+19.9%)	n.s.

Legend: n = number of seedlings; Control = distilled water; As 45 dH = Dynamized arsenic trioxide; W 45 dH = dynamized water; n.d. = not done.; n.s. = not significant. The p-values are related to comparisons with control.

4.2   Main results

The non-parametric tests based on ranks were adopted for the *in vitro* wheat growth model (Betti *et al.* 1997; Brizzi *et al.* 2002 and 2005) and tobacco/TMV interaction (Betti *et al.*, 2003)**.**

Dealing with the first one (wheat growth model) we applied Mann-Whitney rank sum test, considering control and treatment groups as independent. The results are reported in Table 2; it is easy to note that As 45 dH treatment always induced a significant stimulating effect on wheat seedlings. In the second

experiment (Brizzi *et al.,*2002 and 2005) the sample size was much lower because we checked five different potencies simultaneously and, as usual, As 45 dH (reported here) showed the most significant results in wheat growth stimulation**.** We can also observe that the percentage stimulating effect is almost the same (+24.0% and +24.7%, respectively), even if the control mean, is considerably different in the two experiments, due to seasonal effect (the first in winter, the second in summer). Regarding dynamized water (W 45 dH), we did not detect any significant effects in this model, although there was an increase of almost 20% in mean stem length, as reported in Table 2.

As mentioned above, when working with tobacco plants inoculated with TMV, we decided to perform the Wilcoxon test for dependent samples; the results are summarized in Table 3. .It can be seen that both treatments (As 45dH and W 45dH) induced a highly significant mean reduction of necrotic lesion number with respect to the control, although there is an exception for dynamized water in the third experiment. This effect can be related to an improvement of plant resistance due to homeopathic treatment. The significant results obtained with dynamized H_2O suggest that solvent dynamization alone is able to induce effects similar, but weaker, than homeopathic arsenic, as seen in the wheat germination model in Table 1.

Table 3: Tobacco/TMV interaction results: mean number of necrotic lesions and Wilcoxon test significance

Reference	n	Control	As 45 dH		W 45 dH	
		Mean	Mean	p-value	Mean	p-value
Betti et al. (2003b),1st exper.	90	76.6	58.6	< 0.001	68.7	< 0.001
Betti et al. (2003b),2nd exper.	90	118.2	81.4	< 0.001	64.4	< 0.001
Betti et al. (2003b),3rd exper.	90	93.3	85.1	< 0.001	108.8	< 0.001
Overall experiment	270	96.0	75.0	< 0.001	80.6	< 0.001

Legend: n= number of leaf disks; Control= distilled water; As 45dH = Dynamized arsenic trioxide; W 45 dH = dynamized water. The p-values are related to comparisons with control.

5   Checking variability

5.1 Theory and methods

An important analysis was performed to check variability: indeed, we considered it not only as a useful tool for evaluating mean results, but as a marker of primary importance for detecting the effects of homeopathic high dilutions. In fact, we

observed a repeated and not negligible decrease in variability in treated groups for all the models we considered.

Since the mechanism of action of homeopathic treatments is not clear at all, we tried to deepen our analysis by splitting variability into its two components, distinguishing variability "within" experiments from variability "between" experiments and making statistical comparisons for both components.

If we indicate the number of experiments with q, the k-th observation of the i-th experiment with y_{ik}, the sample size and mean value of the i-th experiment with n_i and m_i, respectively, n being the overall sample size and m the global mean value, we have:

$$SD \text{ (global)} = \sqrt{\frac{\sum_{i=1}^{q}\sum_{k=1}^{ni}(y_{ik}-m)^2}{n}} \tag{5.1}$$

$$SD_W \text{ (within experiments)} = \sqrt{\frac{\sum_{i=1}^{q}\sum_{k=1}^{ni}(y_{ik}-m_i)^2}{n}} \tag{5.2}$$

$$SD_B \text{ (between experiments)} = \sqrt{\frac{\sum_{i=1}^{q}(m_i-m)^2 n_i}{n}} \tag{5.3}$$

We adopted standard deviation as a marker of variability, because it has the same unit of measurement of data, but we have to consider that, due to the square root effect, the sum $SD_W + SD_B$ can no longer be exactly equal to SD.

5.2 Main results

Looking at Table 5, we can easily observe that variability between experiments (expressed by SD_B) in treated groups is smaller than in the control, in all plant models we considered. Sometimes the difference is really strong, reaching 30% or even 50%. Moreover, variability within experiments (expressed by SD_W) is also generally smaller with only two exceptions. The resulting overall standard deviation (see Table 4) is thus a decrease in treated groups, with just one exception with an almost immaterial increase. It is worth to consider that this decrease in variability, especially in the component "between experiments", can not be a simple consequence of a reduced mean level of data magnitude, since it has been observed also in wheat growth experiments, where the mean value of the treated

group was higher than in the control group (see Table 2). These findings seem to suggest that homeopathic treatments may have a significant effect on variability.

Table 4: Global standard deviation observed in different plant models (treatment vs. control).

Plant models		Standard Deviation		
Wheat germination:	Stress	Control	As 45 dH	Diff. %
Betti et al. (1994)	No	1.35	0.89	- 34.3
Brizzi et al. (2000), 1st exper.	No	1.33	1.03	- 22.8
Brizzi et al. (2000), 2nd exper.	No	1.58	1.27	- 19.8
Brizzi et al. (2000), 1st exper.	Yes	2.37	2.14	- 9.7
Brizzi et al. (2000), 2nd exper.	Yes	3.50	2.41	- 31.1
Brizzi et al. (2009)	Yes	2.47	1.86	- 24.8
Wheat growth:				
Betti et al. (1997)	Yes	1.98	2.03	+ 2.4
Brizzi et al. (2002, 2005)	Yes	3.63	2.57	- 29.3
Tobacco/TMV:				
Betti et al. (2003 b)	Yes	59.99	51.18	- 14.7

Table 5: Standard deviation components observed in different plant models (W = within experiments, B = between experiments).

Plant models	n.of	Std. Dev. (W)			Std. Dev. (B)		
	exp	Control	As 45dH	Diff.%	Control	As 45dH	Diff.%
Wheat germination:							
Betti et al. (1994)	16	1.04	0.60	- 42.3	0.86	0.65	- 24.3
Brizzi et al. (2000) 1st exper.	12	1.12	0.86	- 23.4	0.71	0.56	- 21.4
Brizzi et al. (2000) 2nd exper.	8	1.32	0.94	- 28.6	0.87	0.85	- 2.4
Brizzi et al. (2000) 1st exper.	12	1.32	1.46	+10.8	1.97	1.56	- 20.6
Brizzi et al. (2000) 2nd exper.	12	3.09	2.27	- 26.5	1.64	0.81	- 50.8
Brizzi et al. (2009)	8	2.26	1.62	- 28.2	1.01	0.91	- 10.0
Wheat growth:							
Betti et al. (1997)	8	1.79	1.91	+ 7.0	0.85	0.67	- 21.3
Brizzi et al. (2002, 2005)	3	3.48	2.46	- 29.4	1.03	0.74	- 28.0
Tobacco/TMV:							
Betti et al. (2003 b)	3	57.52	49.83	- 13.4	17.06	11.69	- 31.5

6 Concluding remarks

First of all, it seems worth recalling that, when working with such a particular field of research like homeopathy, it is necessary to apply a thorough statistical analysis, as well as adapting statistical tools to specific problems that we have to face. After dealing for many years with all the statistical methods described above, applied to different plant models, we can make some general remarks:

- The Poisson model fits our wheat germination data very well, and this allows us to make some parametric inferences without worrying about population distribution. From a biological point of view, it helped us to detect the repeatedly significant stimulating effect of As 45dH with respect to the control, much more evident when seeds were previously stressed. Moreover, the statistical analysis based on Poisson inference allowed us to demonstrate that the dynamization process is a factor of primary importance for the efficacy of homeopathic treatment: indeed, dynamization of water itself induces some significant effects when compared with the control.
- Since our plant models often allowed us to work with large samples, we were able to apply non-parametric tests based on ranks (Mann-Whitney and Wilcoxon), having almost the same power as parametric tests (Student's *t*). We thus repeatedly observed highly significant results, confirming the stimulating effect of homeopathic arsenic on wheat seedling growth, as well as the increase of tobacco resistance to TMV.
- The specific interest given to statistical methods and results allowed us to define a new criterion for evaluating the observed results, where variability may be considered as a primary target of homeopathic treatment action (Betti *et al.* 2003 b, Nani *et al.* 2007). The regularity in variability decrease, and markedly in variability "between experiments" is impressive, and a similar effect has also been detected and pointed out by Binder *et al.* (2005). These findings induce us to suggest this double approach (based on mean effect and variability) when carrying out basic research in complementary medicine.

References

[1] Betti, L., Brizzi, M., Nani, D., and Peruzzi, M. (1994): A pilot statistical study with homeopathic potencies of Arsenicum Album in wheat germination as a simple model. *British Homeopathic Journal*, **83**, 195-201.

[2] Betti, L., Brizzi, M., Nani, D., and Peruzzi, M. (1997): Effect of high dilutions of Arsenicum album on wheat seedlings from seed poisoned with the same substance. *British Homeopathic Journal*, **86**, 86-89.

[3] Betti, L., Borghini, F., and Nani, D. (2003a): Plant models for fundamental research in homeopathy. *Homeopathy*, **92**, 129-130.

[4] Betti, L., Lazzarato, L., Trebbi, G., Brizzi, M., Calzoni, G.L., Borghini, F., and Nani, D. (2003b): Effects of homeopathic arsenic on tobacco plant resistance to tobacco mosaic virus. Theoretical suggestions about system variability, based on a large experimental dataset. *Homeopathy*, **92**, 195-202.

[5] Betti, L., Trebbi, G., Nani, D., Majewski, V., Scherr, C., Jäger., T., et al. (2008): Models with plants, microorganisms and viruses for basic research in homeopathy. In Bonamin, L.V. (Ed.): *Signals and Images*. London: Springer, 97-111.

[6] Binder, M., Baumgartner, S., and Thurneysen, A. (2005): The effects of a 45x potency of Arsenicum album on wheat seedling growth – a reproduction trial. *Forschende Komplementärmedizin und Klassische Naturheilkunde*, **12**, 284-291.

[7] Brizzi, M. and Betti, L. (1999): Using statistics for evaluating the effectiveness of homeopathy: analysis of a large collection of data from simple plant models. 3° *Congresso Nazionale della Società Italiana di Biometria*, Roma, Abstracts, 74-76.

[8] Brizzi, M., Nani, D., Peruzzi, M., and Betti, L. (2000): Statistical analysis of the effect of high dilutions of arsenic in a large dataset from a wheat germination model. *British Homeopathic Journal*, **89**, 63-67.

[9] Brizzi, M., Biondi, C., Lazzarato, L., and Betti, L. (2002): Analisi esplorativa dell'effetto di soluzioni ultramolecolari di triossido di arsenico sullo sviluppo vegetativo *in vitro* di plantule di grano. *Statistica*, **LXII**, 353-360.

[10] Brizzi, M., Lazzarato, L., Nani, D., Borghini, F., Peruzzi, M., and Betti, L. (2005): A biostatistical insight into the As_2O_3 high dilution effects on the rate and variability of wheat seedling growth. *Forschende Komplementärmedizin und Klassische Naturheilkunde*, **12**, 277-283.

[11] Brizzi, M., Elia, V., Trebbi, G., Nani, D., Peruzzi, M., and Betti, L. (2009): The efficacy of ultramolecular aqueous dilutions on a wheat germination model as a function of heat and aging time. *eCAM, Evidence-based Complementary and Alternative medicine.* doi:10.1093/ecam/nep217, 1-11.

[12] Conover, W.J. (1980): *Practical Nonparametric Statistics*. Second Edition. Hoboken: John Wiley and Sons.

[13] Elia, V., Napoli, E., and, Germano R. (2007): The "Memory of water": an almost deciphered enigma. Dissipative structures in extremely diluted aqueous solutions. *Homeopathy*, **96**, 163-169.

[14] Hoaglin, D.C. (1980): A Poissonness plot. *The American Statistician*, **34**, 146-149.

[15] Lüdtke, R. and Rutten, A.L.B. (2008): The conclusions of the effectiveness of homeopathy highly depend on the set of analyzed trials. *Journal of Clinical Epidemiology*, **61**, 1197-1204.

[16] Nani, D., Brizzi, M., Lazzarato, L., and Betti, L. (2007): The role of variability in evaluating ultra high dilutions effects: considerations based on plant model experiments. *Forschende Komplementärmedizin und Klassische Naturheilkunde*, **14**, 301-305.

[17] Sachs, L. (1984): *Applied Statistics. A Handbook of Techniques*. New York: Springer Verlag.

[18] Shang, A., Huwiler-Müntener, K., Nartey, L., Juni, P., Dorig, S., Sterne, J.A., *et al.* (2005): Are the clinical effects of homeopathy placebo effects? Comparative study of placebo-controlled trials of homeopathy and allopathy. *Lancet*, **366**, 726-732.

[19] Stephens, M.A. (1974): EDF statistics for Goodness-of-fit and some comparisons. Journal *of the American Statistical Association*, **69**, 730-737.