Journal of Ayurveda and Integrative Medicine xxx (2018) 1–6

Contents lists available at ScienceDirect

Journal of Ayurveda and Integrative Medicine

journal homepage: http://elsevier.com/locate/jaim

Bias and misleading concepts in an *Arnica* research study. Comments to improve experimental Homeopathy

Salvatore Chirumbolo [a, *], Geir Bjørklund [b]

[a] Department of Neurological and Movement Sciences, University of Verona, Italy
[b] Council for Nutritional and Environmental Medicine (CONEM), Mo i Rana, Norway

ARTICLE INFO

Article history:
Received 9 December 2016
Received in revised form
23 January 2017
Accepted 30 January 2017
Available online xxx

ABSTRACT

Basic experimental models in Homeopathy are of major interest because they could get insightful data about the ability of high dilutions to work in a biological system. Due to the extreme difficulty in the highlighting any possible effect and trusting its reliability, methods should be particularly stringent and highly standardized. Confounders, handling process, pre-analytical errors, misleading statistics and misinterpretations may lead to experimental biases. This article tries to elucidate those factors causing bias, taking into account some recent reported evidence in the field.

1. Background

Marzotto et al. reported that a plant extract from *Arnica montana* L. from Boiron Laboratoires (Lyon, France) in 30% v/v EtOH/water contained 36.0 mg/100 ml of sesquiterpene lactones, namely 1.05×10^{-3} M. The 1:100 preparation named 1c, was used as the starting solution for a series of further 1:100 dilutions in 30% v/v EtOH/water, which showed an effect on the expression of some extracellular matrix genes when tested on IL-4 polarized THP-1 cells [1]. The paper is particularly interesting but raised fundamental concerns about the experimental setting in basic Homeopathy, which is the objective of this article.

First, in order to calculate the molarity of sesquiterpene lactones in the alcoholic preparation, the authors referred to Staneva et al. who identified at least eight components in an *Arnica* extract related to helenalin and dihydrohelenalin by ^{1}H NMR spectroscopy and assumed an average molar mass for dihydrohelenalin-derived compounds of 340.41 [2]. The calculation evaluated by Marzotto et al. which does not rely on any reported chromatographic data, would be an approximation to the estimation done by Staneva et al. with ^{1}H NMR. Staneva et al. reported possible errors in the quantitative analysis performed by using only the average molar mass, particularly for compounds such as methacryloyl-helenalin and assessing that the molar mass calculated by summing the molar weights of single lactones, particularly for isobutyryl-helenalin, 6-O-(2-methylbutyryl)-helenalin, 2-methylbutyryl dihydrohelenalin, which cannot be separately evaluated, was lower [2]. The molar estimation calculated by Marzotto et al. in *Arnica* 1c should refer to the main sesquiterpene lactones present in *A. montana* (erroneously reported as *Arnica m.*), i.e. helenalin and 11α,13-dihydrohelenalin esters, giving the reported theoretical molarity [1,2]. The *A. montana* 2c made in 30% v/v EtOH/water should therefore contain 51.43 mM EtOH and 10.5 nM sesquiterpene lactones. If considering helenalin and 11α,13-dihydrohelenalin as the major compounds from *A. montana* in the extract, the authors showed an effect using doses at least three orders of magnitude lower, than those ones previously reported as effective on *in vitro* immune cells [1,3–5]. If true, this interesting result raises the conundrum of the activity associated with further dilutions, e.g. *Arnica* 5c, as this preparation should be made by 51.43 mM EtOH and 0.0001 fM sesquiterpene lactones, with a ratio EtOH/lactones $= 5 \times 10^{14}$ to 1, a circumstance for which it is very difficult to exclude the molar activity of ethanol with respect to the negligible one of lactones. The same UV–VIS performed by the authors on *Arnica* 1c shows clearly solely the UV absorbance of ethanol, at 205 nm for an $A_{1cm} > 1.0$ at its lowest ε_M value [1]. Therefore, the molar fraction of the chemical components in an *A. montana* L. extract, would suggest that ethanol is the only chemical bioactive species aside from water, which should be present in the centesimal dilutions.

* Corresponding author.
E-mail address: salvatore.chirumbolo@univr.it (S. Chirumbolo).
Peer review under responsibility of Transdisciplinary University, Bangalore.

http://dx.doi.org/10.1016/j.jaim.2017.01.014

Please cite this article in press as: Chirumbolo S, Bjørklund G, Bias and misleading concepts in an *Arnica* research study. Comments to improve experimental Homeopathy, J Ayurveda Integr Med (2018), http://dx.doi.org/10.1016/j.jaim.2017.01.014

2. Ethanol as a confounder and controls performance

The role of ethanol should be better highlighted even though it is difficult to believe that the dilutions may work due to its existence. Ethanol was used in several experimental papers using homeopathic dilutions [6–8], raising comments elsewhere [9–11]. Verma et al. reported recently that ethanol is able to induce the release of nanosized membrane extracellular vesicles able to induce macrophage activation [12]. There is no doubt that ethanol has a chemical activity in those systems where the molar mass of the active principle is absolutely negligible [10,13,14]. However, the most frequent criticism to this comment is that ethanol is present both in controls and in herbal dilutions (cases) and hence, this solvent could not be considered a statistical confounding [15]. If ethanol is present, at the same concentration, both in dilutions and in controls, its confounding effect should be negligible or even null. However, this is true only if both controls and cases are processed in a double blind fashion and are prepared with the same procedure and handling in a high stringency condition. Pre-analytical biases may occur in this case. Batch-derived biases were even reported for gene microarray, particularly in pooling the RNA samples [16,17] and therefore, any difference in the handling, storage and treatment of the ethanol batches of dilutions may interfere and affect the reliability of the results. We must admit that, from a chemical point of view and based on the issues addressed above, a control 30% EtOH/water is perfectly similar to, e.g. an *Arnica* 15c into 30% ETOH/water, because both systems are practically made by only ethanol and water mixed together, due to the negligible or even null amount of sesquiterpene lactones (SLs) in the 15 [1]. Therefore, researchers are most probably comparing two controls with each other, a “control” (A) vs a “dilution” (B). Furthermore, if B undergoes 0.22 μm filtration and A not, as reported [1], if B comes from a batch stored for 12 months while A is a fresh preparation, if B comes from the serial dilutions of previous 4–5 iterative dilutions while A is made from only the previous dilution, and if a blind setting is not considered, differences in the batch of chemical systems that are practically controls at all, may generate bias in the statistics of the outputs and misunderstanding in the conclusive remarks on the reported evidence. Confounders are therefore occurring in the chemical activity of the solvent (ethanol) and in control biases.

3. A statistical evaluation can shed a light on possible biases

Controls should have a highly homogeneous distribution of their inner variance. Previous comments on data variability in experimental Homeopathy showed that even the distribution of standard error of mean (SEM) may lead to statistical significance, due to alpha error in a H_0 null hypothesis [9].

If controls have no homoscedasticity in their data distribution, then a control may bear a “biological” effect due to the existence of a chemical confounder. The herbal dilution with an EtOH/active principle ratio $> 10^{10}$ to 1, is practically a control, where the only chemical active species is the alcohol, as SLs are negligible or practically lacking [1]. Ethanol has a specific activity on gene expression and on differentially expressed genes (DEGs) and may cause bias in the estimation of p values, particularly if performed with an approach, such as Friedman test, which has the very low power of the sign test [14,18]. In this case, effects can be related to ethanol as the main confounder of the dilutions [1,10,15].

To give a possible example of this issue, in a recent paper, statistics was performed using a Friedman sign test, which is less powerful than other non-parametric rank tests, such as the Wilcoxon–Mann Withney test [1,18]. This evidence resembles previous reported data, with RT-PCR [6]. According to the authors, any dilution in 30% v/v EtOH/water was able to change DEG patterns, with $p < 0.05$ in a Wilcoxon test [1]. An evaluation performed on data from Supplementary Tables in the paper [1], using a non-parametric Wilcoxon–Mann Withney rank test, gave the results showed in Table 1. The simple matching of any dilution, from 3c to 15c, vs the control mean (averaging 5 separate experiments) using the RPKM values, gave the following results (bold character = non-significant, i.e. $p > 0.05$ outputs).

a) [*A. montana* 2c] $p = 0.01108$; [*A. montana* 3c] $p = 0.01242$; [*A. montana* 5c] $p = 0.0226$; [*A. montana* 9c] $p = 0.01684$; [*A. montana* 15c] $p = 0.0477$
b) [*A. montana* 2c] $p =$ **0.35238**; [*A. montana* 3c] $p =$ **0.09102**; [*A. montana* 5c] $p =$ **0.37346**; [*A. montana* 9c] $p =$ **0.22628**; [*A. montana* 15c] $p =$ **0.65994**, (bold character $p > 0.05$, i.e. not significant), assessing therefore a circumstance that can be retrieved also from Tables 1 and 2, where dilutions are matched to controls of each single experiment. Data suggest that the variance distribution within the control series is not homogeneously dispersed and give possible biases in the interpretation of the presumptive working of homeopathic dilutions. To ascertain control homoscedasticity, a Bartlett's test should be accomplished. The Bartlett's test on the control distribution showed that this variability was highly significant ($p < 0.0001$, $\chi^2 = 409.19452$). The overall RPKM evaluation of the signed rank comparison between all averaged controls and means for each tested dilution, gave the following statistics:
c) [*A. montana* 2c] $p =$ **0.13622**; [*A. montana* 3c] $p =$ **0.23404**; [*A. montana* 5c] $p =$ **0.21498**; [*A. montana* 9c] $p =$ **0.21499**; [*A. montana* 15c] $p =$ **0.17702**, which should suggest the existence of a possible bias in the distribution used to evaluate the dilution activity on THP-1 cells, as these comparisons would indicate the complete absence of effects on the gene expression of macrophages by *A. montana* alcoholic extracts. This evidence seems to contradict the conclusive remark forwarded by the authors about the activity of *Arnica* [1]. Goodness of fit test, performed with a Shapiro–Wilk test and a Lilliefors-van Soers test assessed that any distribution was non parametric. The number of outliers in a Rosner's extreme studentized deviate test ($p < 0.00001$, $\geq$10 out of values) was 2.25 higher for controls than for any test solutions.

Apparently, the authors did not seem to have addressed this concern. A possible reason is the following. The false discovery rate (FDR) approach has been standardized for barcoded cDNA of samples in a RNA-seq library and sequencing [19] and actually sample pooling yet shows many critical aspects, so that the increase of replicate samples has been suggested as the best choice [20]. Particularly, when negligible concentrations of active principle are challenged with an *in vitro* model of gene expression, a throughput RNA sequencing method should encompass stringent criteria for the statistical evaluation of DEGs In this context, even the concordance of an NGS with a gene microarray in the case of a genome-wide array of differential gene expression, is affected by the treatment effect size, depending by the transcript abundance and the biological complexity of the different modes of action of the tested chemicals, their dosage and ability to interact with genes [21]. Interestingly, ethanol, at the concentration 51.43 mM, i.e. 0.03% v/v into water, is particularly active on a biological system. Ethanol may cause mitochondrial injury [22] and even mitochondrial DNA damage [23] and in doses as low as 50 mM of ethanol is able to cause mitochondria damage, oxidative stress and apoptosis in several cell models [24–27], as 50 mM ethanol may cause 2.03% apoptosis in cardiomyocytes and 4.32% apoptosis in 24 h treated

Please cite this article in press as: Chirumbolo S, Bjørklund G, Bias and misleading concepts in an *Arnica* research study. Comments to improve experimental Homeopathy, J Ayurveda Integr Med (2018), http://dx.doi.org/10.1016/j.jaim.2017.01.014