

Problems of Evidence Based Medicine in Judging Naturopathic treatments

ANME Conference, Frankfurt; November 11, 2006

The question in how far evidence based medicine (EBM) is applicable to the different methods of naturopathy is controversy discussed. In the first place I would like to highlight the fact that the disputes about “EBM and naturopathic treatments” are held on a quite low scientific level. This has changed in recent years on the part of naturopathic treatments, though not voluntarily. But on the orthodox medicine’s side, the scientific level is still unfortunately pretty low. The article on homeopathy in the LANCET has shown this very well (10). This should not hinder advocates and users of naturopathic treatments from further improving their own scientific knowledge. I would like to advise people strongly not to use the complaisance of statistics for economic and other interests as an argument. Of course what I call the “human factor”, the manipulation of results, is a grave problem in medicine. Particularly the pharmaceutical industry has various means to wield influence which leads to a situation far away from of every evidence based state. However, the supporters of EBM are trying intensely to control this problem. Although the “human factor” is a big problem, it is not the essential criterion for criticism of EBM when it comes to naturopathic treatments. I would like to demonstrate this with some examples. Each example illustrates deeper rooted theoretic-scientific difficulties, but due to a lack of time I cannot go into greater details.

1. Those who cure are not necessarily wrong

Often patients come to us and hand us a bag full of evidence proven drugs; drugs that they had taken for years without an improvement of their health state. If now after one year of naturopathic treatments patients are cured without taking any medication anymore, then this is not a proof of the efficiency of naturopathic treatments. It is, however, a proof that there is problem with the usual efficacy control.

2. The impossibility of individual statements

Statisticians report particular difficulties that accompany randomised trials. Simpson’s paradox is probably the funniest of them. It states that the probabilities that are true for two individual populations do not necessarily have to be true for these two populations together (7). This means that something that in a study for women and in another study for men has been proven good, can still be bad for the people altogether (1). Another example: if therapy B turns out to be better than therapy A and if another study shows that therapy C is better than therapy B, then it is a fallacy to believe that therapy C is better

than therapy A (2). Another major problem is that statistics can only make statements on defined populations but not on individual cases. In fact, modern medicine has no means to make individual predictions on scientific bases. Critics of the statistic approach once said, “Big numbers lead to a statistically exact result – but no one knows for whom it is the case. Small numbers lead to a statistically unusable result – but one knows better to whom this applies to. It is hard to say which of these forms of ignorance is more useless (4)”.

3. What might be helpful in the short run, might be harmful in the long run.

I would like to illustrate my next point with some studies.

- in the treatment of epicondylitis it could be shown that an injection of cortisone leads to a significant improvement in the short run. On a long term basis, though, this therapy is remarkably inferior to physiotherapy or the simple waiting (22).
- A study concerning Baby-Walkers showed that those infants adopt an upright posture earlier, but learn to stand and walk much later (14).
- Babies treated with antibiotics in the first six months of their lives have a greater risk of suffering from allergies later on (16). This statement is not undisputed, however (5).
- Preterm babies who receive enriched food are more likely to develop an insulin resistance than those who are fed with breast milk or normal baby formula milk. That means that the quick reaching of a reference level is not healthier (21).

If the principle of “helps in the short run, harms in the long run” occurs more often, then those studies only spanning a couple of months (most of them) and proving the efficiency of a therapy, were nothing less than a manual to harm patients.

4. The more disturbing the context, the higher the specific efficiency

The following phenomenon shall be exemplified by balneology. Those of you who have already been at old spas and got some treatments know the healing effect of a beautiful architecture and its corresponding ambience. It conveys the feeling of internal dignity. Unfortunately, nowadays balneological treatments are given in merely functional buildings that have no charm at all. In such settings specific aspects are examined in an evidence-based manner, for example, which effect certain exercises have on the mortality of patients suffering from a left-ventricular systolic dysfunction (12). What is the use of

these studies? Are they indeed independent from the external frame? And what should we think of a medicine that considers beauty and dignity at best as a placebo effect? What you hear is not the sigh of an aesthete. We are rather dealing with an important scientific problem. In an experiment with hamsters it could be proven that wounds heal quicker if the hamsters were left in their normal social surroundings. When put into a glass tube, their wounds initially got worse (11). That is, if in such a setting the efficiency of a healing substance were examined, it could turn out in the first instance as a placebo, and in the second instance as a specific effective medication. This means that a substance will seem to be more effective, the more harmful the therapeutic framework is. The contradiction between hospital-based research (glass tube) and a general practitioner's treatments is well known, but the consequences have not been taken seriously into account. This problem is of even higher importance in naturopathic treatments, as most of these treatments create a positive therapeutical ambience.

5. Adaptability versus rigidity

Changing contexts play a role in a different respect as well. Beta blockers are (evidence based) good for several illnesses ranging from hypertension to glaucomas and states of agitation. A side effect is the lower ability to adapt to heat and a heat stroke is more likely to happen (8). If climate change leads to hotter summers, then the positive effect found in studies would be gone instantaneously. This implies that a well measurable effect which increases rigidity could be devastating if circumstances change. A hardly measurable effect, on the contrary, leading to an improvement of one's adaptability could be of great use under certain circumstances. So, what might be helpful in one context might be harmful in a different context. Randomised studies cannot differentiate between rigidity and adaptability.

6. *Learning in higher order*

I would like to present a logical instrument that is apt for illustrating this process more accurately. It is the different orders of learning as introduced by Bateson (3).

- Learning of the order 0 exists if a unity shows only little change or none at all in its reaction pattern after having been exposed to repeating stimuli. This form of learning can be found in simple automatic regulatory circuits or if an organism is overstimulated or if a stimulus answer is determined structurally.

- Learning of the 1st order is what we typically consider learning. Learning a language belongs to this category.
- Learning of the 2nd order is learning to learn. When learning a foreign language, one learns to learn other languages more easily. Learning of this category can be found everywhere, but it is not necessarily always positive. If for example a person has a bad experience when learning a language, he or she might be reluctant to learn another one. Then learning another language results more difficult or even impossible. The problem is that learning of the 2nd order cannot be discovered through simple measurements. Rather a row of measurements are needed and a corresponding theory.
- Learning of the 3rd order is characterised by basic changes that go beyond what is actually supposed to be learned. A change from a linear causal-oriented thinking towards a systemic solution-oriented thinking is learning of the 3rd order. It is not about contents but rather about the basic evaluation of contents.
- Learning of the 4th order would be a change of the learning structure and includes morphological and genetic influences.

In the medical context for example, the acquisition of a specific immunity is learning of the 1st order. If a certain infection induces a better or worse resistance against other infections, then we are dealing with learning of the 2nd order. If a chronic disease deteriorates or ameliorates through such an infection, then it is learning of the 3rd order. The last category would imply the epigenetic consequences of such an infection. The basic problem is that current medicine only knows learning of the order 0 and the 1st order. The learning of higher orders is eliminated from the observation framework as an unspecific effect although it is absolutely specific (15).

Randomized studies concerning naturopathic treatment

Randomized studies examine the effectiveness of a specific intervention in a defined setting compared to a sham intervention. This setting, though, does not coincide with the setting of the naturopathic treatment. The keyword is individualization. I would like to illustrate that by the following example.

Even a simple massage has to be individually adapted for every patient, otherwise the massage can not be successful. What is good for one person can be harmful for another. So how to evaluate a study which compares the effect of a standardized physiotherapy (massage,

warmth, cold, etc.) with the simple, standardized instruction to stay active? The result that physiotherapy is not superior to this simple advice, might not be of great relevance. What is important, however, is, that both treatments (standardized therapy and standardized advice) are bad medicine. Considering this form of effectiveness control for naturopathic treatments, one could set the following rule-of-thumb: *The better a study in its methodology, the worse the practiced medicine.*

Basically the common studies measure a form of naturopathic therapy which does not comply with real conditions. The concept of a specific causative factor in a specific situation is, in most cases, not suitable for methods which mainly are aiming at an unspecific activation of self-regulation.

Furthermore, formal logic says that studies can only lead to statements about the things they examine. The above described physiotherapy study thus only proves, that this form of standardized therapy is not superior to a simple advice. Also when a potentiated house dust mite does not show a better result than a placebo in a study on house-dust allergy (17), this does not prove anything about potentiated substances, nor about homeopathy. In the same way, a study which does not show a positive effect of acetylsalicylic acid on breast cancer (9), cannot evaluate the effectiveness of acetylsalicylic acid as such, and the effectiveness of pharmacotherapy in general. Nevertheless wrong conclusions like that are the order. In logics this is called “violation of the logical types” (19). If one class (therapy) is put into the same category with an element of the class (specific intervention), then a paradox situation is created, which is no longer able to give a reliable statement. The idea of examining a whole therapeutic approach on the basis of a study is, therefore, a major problem.

Of course, there have been attempts to adapt studies to the naturopathic treatment setting. The Munich headache-study (23), for example, fulfilled the preconditions of statisticians and homeopaths a priori. Patients with long-term headaches (average duration of illness 23 years) were treated in an individualized way. The result was that the individually designed homeopathic remedy did not show a better effect than the placebo. With the wisdom of hindsight, the homeopathic camp, like always when a study does not show the desired results, raised a number of objections. The study was too short; the cases were too difficult for the simple scheme of the study, etc. In this context it does not matter if those arguments are valid or not. What is important is the fact that the homeopathic treatment as such (remedy and placebo), was effective for almost a fourth of the patients. This result is insofar interesting, as a number of studies on acupuncture came to a similar result (6, 18, 20): Both therapies, remedy and placebo, are effective, but the actual remedy is not or only slightly superior to the

placebo. That means, homeopathy and acupuncture are effective forms of therapy, which are able to control long-term pain in a cost-efficient way and with little side effects. But they don't do that in a specific way.

According to the common logic, the medical standard treatment should be no longer supported by the National Health Insurance as its results are worse than those of an unspecified placebo therapy! Of course nobody would take such a demand seriously. And it is absurd, as it is based on a wrong logic. But also the reversed conclusion, that naturopathic treatment is ineffective, is impermissible. The phenomenon that therapies are effective although they are ineffective is typical for a flawed logic by violating logical types, when the class (method) is mixed up with the element of the class (specific therapy). That means that the results of these studies prove neither the effectiveness nor the ineffectiveness of naturopathic treatments. At best they prove that this type of effectiveness measurement, due to logical reasons, is not appropriate to attain a solid result.

Conclusion

Randomized studies are known as the gold standard of EBM. But in spite of the enormous triumphs of EBM, there are a huge number of second thoughts pointing to the limitedness of this approach. With this type of effectiveness measurement, basic topics like adaptation, individual transferability, and evaluation of complex and/or long-term processes can only be evaluated inappropriately.

Randomized studies are the appropriate means for a medicine mainly aiming at the eradication of symptoms. For regulative processes in which the appropriate stimulus for an individual is chosen, the main procedure of EBM is not very useful. One reason is that in so called network pathologies the same stimulus can cause either amelioration or deterioration. The question which remedy or procedure is more effective in headaches is not so important in this context. This completely different understanding of health and diseases can not only be found with naturopathic treatment. The more recent research studies about immunology point at the fact that the conventional understanding of disease based on simple patho-physiological descriptions or even on the counting of criteria, is an end-of-range model. What kind of design the effectiveness control in such a regulative health model (in contrast to a patho-physiological disease model) might have is hard to say. In my opinion it should be a dynamic procedure that can shape transformation processes. If one is to continue to call such a type of effectiveness control, which completely differs from toady's type of control, EBM, depends

on one's personal taste. I personally rather tend to not burden a modern understanding of medicine with old terms.

Bibliography

- (1) Baker SG, Kramer BS (2002): *The transitive fallacy for randomized trials: If A bests B and B bests C in separate trials, is A better than C?* BMC Medical Research Methodology 2002 2:13, available <http://www.biomedcentral.com/1471-2288/2/13>
- (2) Baker SG, Kramer BS (2002a): *Good for Women, Good for Men, Bad for People: Simpson's Paradox and the Importance of Sex-Specific Analysis in Observational Studies.* Journal of Women's Health & Gender-Based Medicine 10; 9: 867 - 872
- (3) Bateson G (1990): *Ökologie des Geistes*, Suhrkamp, Frankfurt, S: 362-399
- (4) Beck-Bornholdt HP, Dubben HH: *Der Schein der Weisen.* Reinbek bei Hamburg: Rowohlt 2003
- (5) Benn CS, Melbye M, Wohlfahrt J, Björkstén B, Aaby P (2004): *Cohort study of sibling effect, infectious diseases, and risk of atopic dermatitis during first 18 months of life*, BMJ 328:1223
- (6) Berman BM, Lao L, Langenberg P, Lee WL, Gilpin AM, Hochberg MC (2004): *Effectiveness of Acupuncture as Adjunctive Therapy in Osteoarthritis of the Knee*, Annals of Internal Medicine 141; 12: 901-910
- (7) Bogomolny A (1996-2003): *Mediant fractions*, available <http://www.cut-the-knot.org/blue/Mediant.shtml#simpson>
- (8) Bouchama A, Knochel JP (2002): *Heat stroke*, N Engl J Med 346: 1978 - 1988
- (9) Cook NR, Lee IM, Gaziano JM, et. al (2005): *Low-Dose Aspirin in the Primary Prevention of Cancer: The Women's Health Study: A Randomized Controlled Trial*, JAMA 294: 47-55
- (10) Dellmour F (2006): *Klinische Studien und Metaanalysen in der Homöopathie*, Deutsche Zeitschrift für klinische Forschung 5/6: 52-60, available http://www.dzmf.de/heft/2006_05-06/12.pdf
- (11) Detillion CE, Craft TKS, Ghasper ER, Prendergast BJ, DeVries AC: (2004) *Social facilitation of wound healing*, Psychoneuroendocrinology 29; 8: 1004-1011
- (12) ExTraMATCH Collaborative (2004): *Exercise training meta-analysis of trials in patients with chronic heart failure (ExTraMATCH)*, BMJ 328:189
- (13) Frost H, Lamb SE, Doll HA, Carver PT, Stewart-Brown S (2004): *Randomised controlled trial of physiotherapy compared with advice for low back pain*, BMJ 329:708
- (14) Garrett M, McElroy AM, Staines A. *Locomotor milestones and babywalkers: cross sectional study.* BMJ 2002;324:1494
- (15) Ivanovas G, Tomaras V, Paritsis (2007): *Human adaptation and conscious purpose in contemporary medicine*, Kybernetes, to be published 2007
- (16) Johnson CC, Ownby DR, Alford SH, Havstad SL, Williams LK, Zoratti EM, Peterson EL, Joseph CL (2005): *Antibiotic exposure in early infancy and risk for childhood atopy*, J Allergy Clin Immunol. 115; 6:1218-24.
- (17) Lewith GT, Watkins AD, Hyland ME, et. al. (2002): *Use of ultramolecular potencies of allergen to treat asthmatic people allergic to house dust mite: double blind randomised controlled clinical trial.* BMJ 324:520
- (18) Linde K, Streng A, Jürgens S, Hoppe A, et. al.(2005): *Acupuncture for Patients With Migraine*, JAMA. 293:2118-2125
- (19) Russell, Bertrand (1930): *Introduction to Mathematical Philosophy.* Allen&Unwin, London
- (20) Scharf HP, Mansmann U, Streitberger K, Witte S, Krämer J, Maier C, Trampisch HJ, Victor N, PhD (2006): *Acupuncture and Knee Osteoarthritis*, Annals of Internal Medicine 145; 1: 12-20

- (21) Singhai A, Fewtrell M, Cole TJ, Lucas A (2003): *Low nutrient intake and early growth for later insulin resistance in adolescents born term. Lancet* 361: 1089-97
- (22) Smidt N, van der Windt DAWM, Assendelft WJJ, Devillé WLJM, Korthals-de Bos IBC, Bouter LM (2002): *Corticosteroid injections, physiotherapy, or a wait-and-see policy for lateral epicondylitis: a randomised controlled trial. Lancet* 2002; 359: 657-62
- (23) Walach H, Haeusler W, Lowes T, et. al. (1997): *Classical homeopathic treatment of chronic headaches, Cephalalgia* 1997; 17: 119-126