A short methodology for Beginners
Part 5: Empiricism and theory (1)


Prof. Dr. Harald Walach

In this series of (blog-)articles I'll gradually develop an online tutorial on the key elements of methodology which can serve as a reference for physicians, those interested in medicine,, journalists, methodical laypeople, and students. In the end these entries will comprise a small study compendium. Contributions from others are welcome.

Harald Walach

Part 5: On the relationship between empiricism and theory (1) or: Can we be converted by data?

In the last chapter, "EBM in action," we saw that: The current guidelines of the American Society of Anesthesiology recommend non-steroidal anti-inflammatory drugs (NSAIDs) for the treatment of chronic back pain, as a simple analgesic among other drugs. This recommendation is based on five studies, two of which relate to pain medications taken off the market due their side effects. Two of the remaining three studies were so small and short that they are not very meaningful. This guideline also ignored an authoritative overview of 53 relevant studies published in 2000 which comes to the conclusion that painkillers are useless for the treatment of chronic back pain.

I have not chosen this example because I wanted to denigrate anyone, but because I wanted to use it as my research and demonstration assignment to check transferability of clinical trial data to practice populations. Ultimately, this question is irrelevant and can not be answered, because another question came into the spotlight: How is it possible that a scientific society published such a guideline despite the presence of overwhelming negative data received from the widely acclaimed "evidence" of conventional medical research? How can it be that scientific data – retrieved from scientific evidence-based medicine - are so blatantly ignored?

There are at least three reasons:

1. **The clinical experience of the physician is also part of EBM**. The clinical experience is included in the EBM evaluation through the survey of the technical panel members. And this included experience can be quite positive, even though scientific data shows a different picture. Why? Because the placebo effect plays an

enormous role, and because you can achieve surprisingly good results with the mobilization of hope, expectation, relaxation and through simple conditioning, meaning learning experiences from previous treatments. It's not too bad, in fact. I've always been of the opinion that the best therapist is the one who mobilizes the self-healing effects in patients. However, in the case of NSAIDs these effects are accompanied by an enormously high side effect potential. Therefore, it would probably be not only wiser, but also more ethical, to use therapies that have very low, sometimes perhaps even no specific effects, but very high placebo effect, at least among the patients who believe in this therapy. These mainly include methods of complementary medicine, acupuncture, homeopathy, radionics, bioresonance and so on. So, lets spell out the reason why NSAIDs are still in the guidelines for chronic back pain treatment: in the hands of the specialists surveyed NSAIDs produced high placebo effects, presumably because those very specialists believe in their effectiveness.

2. **There is a definite bias** . "Bias" is a statistician's expression for "distortion". So here is a distortion of perception resulting in a more positive evaluation of these therapies by the surveyed experts than would be appropriate based on data only. We will not discuss this at great depth at the moment. However, the key to understanding this paradox is found in the so-called conflict of interest. It is well known and has also been frequently condemned by the press that many expert panels publishing the majority of these and similar guidelines consist of researchers who have research grants, honorariums, or even shares from the pharmaceutical companies that manufacture these products. This creates bias. Distortion generates misperception. Misperception creates resistance to the often sad reality.
3. This leads me to the third point and today's topic: **The impact of our previous experience**. We are all generally not as good empiricists as we would like to believe. Instead, we are guided by our preconceptions which can be only changed in a limited way by data. And that's the crux of academia. For although many people, scientists especially, seem to make decisions based on empirical data, in reality the following happens: they formed a certain expectation of how the reality is likely to behave (reasonable expectation) based on their training, their experience, what they heard in their culture and from their colleagues and peers. Everything we experience is incorporated into that world view. Scientific and empirical data are only one form of experience; a very structured and informative experience, but just one among many. The more settled our world view is, meaning made of more prior experiences, the harder it will be for a new experience to change our world view. Most experiences are evaluated as to whether they conform to our expectations or not. If the experiences are consistent with and conform to our expectations, we acknowledge them gratefully with a nod and a good feeling and save them perhaps in the category: "I've always said so" - and go about our business as usual. If the experience or scientific data are inconsistent with our expectation, we have two choices: we ignore them and think " that was an exception" or "outlier, a coincidence". Or we take the experience seriously and need to change our preconception. When do we do that?

> Do we do it at all? I think that we rarely adjust our opinion, too rarely. Because **we are Bayesians** (see below) **by our biological structure.** This means **we tend to confirm our preconception, we will seek information to do exactly that, and we tend to ignore information calling our preconception into question**. This is true in general for everyone. And unfortunately, it is also true all too often among scientists.

I want to argue this point more and explain first why this may be biologically meaningful but scientifically dangerous. I will also briefly explain what a "Bayesian" is, the corresponding statistics will come later. Let's start from the rear:

We are all Bayesians - Or: Why it might be worthwhile to insist on preconceptions

Thomas Bayes and output probabilities

A Bayesian is one who is acting as predicted in the theorem of Thomas Bayes (1702-1761), the Presbyterian priest and mathematician. Bayes also dealt with, among other things, the question of probability. Normally we think of probability as follows: Suppose there are 50 black and 50 white balls in a box. We then ask: What is the probability of drawing a white ball if all the balls are well mixed up? Clearly the answer is “50%”, because there are two options that appear equally often. Bayes turned the tables. He asked himself: Suppose I drew a couple of balls and thus have a little initial information, what do I know about the balls in the box? Bayes hypothesized the *belief* about reality that we can form based on a little initial information, experience or empirical data. And he found out that - it is indeed quite intuitively obvious –the way we assess data, therefore reality, depends on what we know or think we know about it beforehand. If we know little about something, then the prior probability (Bayes terminology) for a particular event is about as large as the probability of a possible contrary event. In this case we achieve a certain security on how we evaluate reality with relatively little empirical work (the so-called "posterior probability"). This is the formalization of the fact that empirical reality and experience change opinions. If we don’t have a strong opinion on an issue we can form our opinion quickly or accept a certain opinion based on a little bit of information. But if we already have a heavily embossed and clear opinion on an issue, it also means that the opposite is very unlikely to us and that we need a lot of empirical data and persuasion before we change our opinion. [1]

Example 1 : If an expert formed the opinion that NSAIDs are useful drugs based on his long training and his reading of countless studies, based on a constant flood of information with news about the usefulness of certain products, and based on the sufficient financial incentives provided by drug manufacturers, then he will give up that view only after very important experiences to the contrary. Such an opposing experience to his views could be, for example, that his wife falls ill due to serious side effects of such a treatment, or perhaps even a series of good studies proving to him the opposite. One study showing the opposite would probably not be enough, it would have to be more than one. A simple review, even if it is systematic, would probably not change anything either, because reviews are also nothing more than a single publication, sometimes flawed, etc.

Example 2 : If a scientist decided for himself that homeopathy can't work at all because it is known that there are no more molecules left in the homeopathic solution, then even whole armies of positive data can not change this opinion. I saw this recently apostrophized at a meeting where Barney Oliver, at the time he was head of research at Hewlett- Packard, wrote: "This is the sort of thing I would not believe, even if it were true." This sounds weird, but is absolutely consistent and Bayesian: once one has a really clear preconception based on other experience, she or he will not give up that opinion so easily.

Our brain is a reality-construction-machine

In my opinion, this firm hold on preconceptions is *biologically pre-formed* and therefore makes sense - within some limits. The question is to what extent and why? This has to do with our neurobiology. [2] *We are born into the world as a fairly blank slate*. While there are a few genetic preconditions, our neural system is at first still not inter-connected, but becomes so in the first weeks, months, years, at an enormous speed and plasticity. This plasticity is preserved, but not with the same dynamic as in the early years of life. Therefore, small children can learn many things much easier than adults - sports, musical instruments, languages, juggling. Young children in China learn Chinese in three years, some adults never learn it. Our experiences shape us and shape the way *our brain* reacts. We now know that it is *very much more constructive then we think* . Raichle, a neuroscientist, once coined the expression "the *brain's dark energy*." This term refers roughly to the following facts: about 90% of all our brain activity is handling internal stimuli, and the brain uses less than 2% of all its energy to process stimuli from outside. In other words, data and information from the outside only modulate what happens in the brain continuously. What happens there? We create a world from our experiences, project this view onto the outside and continue to do this until the comparison with our input raises a red flag. To put it in other words: *We do not perceive the world, but we constantly construct it, based on our previous experiences*. What we consider as our perception of the world is "in reality" (whatever this means by now) our construction which is softly modulated by a regular, but moderate check on reality. We do not work like a digital camera or any other technical device that we have invented that represents reality naively and without preconceptions. We are more like a reality generator checking every now and then that the designed reality doesn't deviate too much from our experience. The organism doesn't have to faithfully reflect reality as a whole for survival, but only incorporate some areas of reality into its awareness and behaviour in a way that secures survival.

Such an approach makes biological sense, because it costs more time and energy to re-create everything new from scratch all over again every second. It is much easier to assume that everything remains as usual - and only that which is different is altered in the internal representation and design.

This is how we work as a biological species. What applies to us as individual beings with complex cognitive apparatus is also true for us as scientists and, therefore, also applies to science, which is a collection of such scientists. It shouldn't surprise us that conventional pain therapists consider NSAIDs to be effective in chronic pain management although it is shown otherwise by current data. Their initial probability for such a statement is very high. This is how they construct their reality. Therefore, before

these conventional therapists could change their mind, they would require a lot of negative data, or a very strong negative experience. And so it does not surprise me, even if it annoys me, that critics of complementary medicine are so obviously blind to data. The initial probability that complementary medicine or homeopathy works is so low in the eyes of their critics that even a giant pile of positive data, which indeed already exists in some areas, would make little, or initially hardly any, difference.

Concretization and examples

You can try it yourself. I found a nifty Internet program that makes the appropriate calculations for you and shows you how conventional statistical results expressed as p-values or error probabilities change preconceptions or must be interpreted differently based on existing preconceptions:

http://www.graphpad.com/quickcalcs/DistMenu.cfm

Go to "interpret a p-value", where you will see an entry menu. This prompts you to define a p-value. Let's say you had a study with a conventional significance level of 0.05. Assume further that the study had proper statistical power, i.e. 90% of input (which I'll explain in more detail in another blog). Now comes the correction for the baseline probability. Let's say you're an expert in a panel and start out with the opinion that drugs are good for the world and pain killers work well, even in chronic cases. Their output probability is also 90%. Click "Calculate" and you will see: a single study of this kind will change your belief that pain medications work in a virtual certainty with a probability of nearly 100% (posterior probability = 0.9939). Suppose you saw a non-significant result in such a situation, it would make the subordinate probability that this was a mistake about 50%. If you assume that the initial probability that painkillers work for chronic back pain is 99%, then a significant result brings virtually 100% security and a negative result will still leave you a 91% margin for interpretation that the result was just a simple mistake.

If we, however, simulate a homeopathy skeptic he or she would assign only a 1% probability that homeopathy works. A significant positive study with 90% statistical power will increase this output probability to 15% only. A negative result, however, will increase his or her 99% certainty that such an effect can't exist to 99.89%, an even closer move towards safety. If a critic believes that homeopathy works with 0.001% certainty, then this positive result will move the probability to 1%.

We see that the output probabilities which we maintain in our heads formed as a result of our past experience, our preconception, our vanity, or whatever, influence even in a formalizable manner how far empirical data can or can't change our preconception and view of reality. **The higher the initial probability, the easier will we accept an empirical result and vice versa**. After all, **we are all Bayesians** . Thus experts can consider NSAIDs as effective for chronic back pain, even though they are not, and homeopathy critics can ignore the existing data (and simultaneously advise others to do so).

Current example: Professor Ernst has said that not even my own negative data on homeopathy can convince me of its ineffectiveness, and therefore I am not a serious scientist. I think that the opposite is true. I have conducted, according to Ernst, one of

the methodologically cleanest clinical trials of homeopathy.[3] This study had a negative result. From this I have drawn the conclusion (that made many homeopaths unhappy) that whatever happens there, has nothing to do with a conventional-causal pharmacology. Otherwise we would see different kind of effects. At the same time I also have a wealth of other experiences to draw from: my own personal experience with homeopathy, many case reports from reliable hands and from the literature. Therefore, I have a different prior probability. From this I have come to the conclusion that something is weird here and the simple hypothesis "all is just placebo" does not suffice. I then performed a couple of drug tests. These were experimental, blinded studies on healthy people. Even the first two pilot studies have revealed interesting data. The main study, consisting of two sub-studies, produced a clear significant effect. A second independent study was also significant [4]. From this I concluded: some homeopathic medicines do, at least sometimes, show symptoms in healthy volunteers that are not placebo. This means according to the logic of science that the statement "Homeopathy is always and under all circumstances equal to placebo" can not be true any longer. And that is it. It is interesting that critics studiously ignore this part of my data. When an editor of “Der Spiegel“ sent me a set of questions about homeopathy per e-mail, I sent her (well aware that she was going to write a damning review on homeopathy) exactly these data with the request to consider them. She didn’t do that. The same editor had publicly announced at a podium weeks earlier that homeopathy belongs to medical history and it was her job to transport it there. This episode is a classic example of preconception and Minister Bayes in action. When something like this happens, data can never be of much help, simply because it would shake a safe world view and that is too uncomfortable. This is also precisely the reason why Edzard Ernst quotes me incorrectly, because the rest of my data does not fit into his theoretical world view. I think that's unfortunate, but understandable. Because people are the way they are. We are all, in general, Bayesians.

Outlook

A good scientist is different from a wannabe scientist or a self-proclaimed science pope through his or her willingness to learn from the data and to reveal and adapt their theory to the experience, as well as through his or her openness to new experiences, even though he or she has their own theory. [5] You might ask how this is possible when we are apparently doomed to live as Bayesians, even from a biological perspective? A very small but effective remedy exists: a systematic forgetting of security and automated processes, a systematic culture of an open mind. One can practice, for example, meditation, which clears and re-aligns the mind and continuously opens us to the mystery of life and its surprises. This is the essence of life, spirituality, and even science. [6] Meditation transforms us from a predetermined Bayesian to open minded humans. If it would be possible to allow any empirical measured treatment a 50% probability from the start, whether it is NSAIDs for chronic back pain or homeopathy, then a single good study with a positive result would be sufficient to convince us that the study reflects reality and a single good study with negative result would do the same. Openness saves enormous resources, and would save us a lot of fights.

Notes:

[1] Bayesian statistics and related thinking is not entirely trivial. Therefore, it is little known although it is actually more accurate than the dominant frequentist statistics. I've found the introductions of the following papers useful: Pamar et al. (2001), Raha (2011) and a very good example by Tressoldi (2011). See also the Wikipedia entries on "Bayesian" and "Bayesian ..." . On the English language site there are some links to further online tutorials, where you can make yourself more knowledgeable.

[2] What I outline here is standard neurobiology. Every textbook contains relevant information. I found Roth (1997) very useful. The central work is Raichle (2006).

[3] See Walach et al. (1997). The corresponding critical criticisms were formulated by Vithoulkas and Oberbaum. My replies in Walach (2002a, b) and a few thoughts on where these data have led me in Walach (2000). You do not have to agree with me on this, but you can see that I've taken the data seriously and drastically changed my preconception.

[4] The first pilot studies were Möllinger et al (2004) and Walach et al (2004). The appropriate follow-up studies were Walach et al (2008) and Möllinger et al. http://content.karger.com/ProdukteDB/produkte.asp?Aktion=ShowPDF&ArtikelNr=209386&Ausgabe=247634&ProduktNr=224242&filename=209
I combined all my drug test data in a book chapter Walach (2009).

[5] Although this is trivial, it is worth emphasizing again and again because it is often forgotten. Especially these days, you can read it in the praises of the new Nobel Laureates over and over again: A good example is Daniel Shechtman, the discoverer of quasicrystals. He saw them in the electron microscope, could not believe his eyes (and experience) and still changed his mind, forming a different opinion from the textbook and his colleagues. He resisted their skepticism for almost 20 years and finally received the Nobel Prize. See http://www.stern.de/wissen/nobelpreis-fuer-den-glauben-ans-unmoegliche-kristall-1735483.html

http://www.sueddeutsche.de/wissen/stockholm-nobelpreis-fuer-chemie-geht-nach-israel-1.1155492.

[6] I have presented this idea in my book spirituality Walach (2011) and a correspondingly smaller chapter in Walach (2008). Soon-perhaps an extra chapter on this topic.

Literature

- Möllinger, H., Schneider, R., Löffel, M., & Walach, H. (2004). A double-blind, randomized, homeopathic pathogenetic trial with healthy persons: Comparing two high potencies. *Forschende Komplementärmedizin und Klassische Naturheilkunde* 11, 274-280.

- Möllinger, H., Schneider, R., & Walach, H. (2009). Homeopathic pathogenetic trials produce symptoms different from placebo. *Forschende Komplementärmedizin* 16, 105-110.

- Parmar, M. K. B., Griffiths, G. O., Spiegelhalter, D. J., Souhami, R. L., Altman, D. G., van der Scheuren, E., et al. (2001). Monitoring of large randomised clinical trials: a new approach with Bayesian methods. *Lancet* 358, 375-381.

- Raha, S. (2011). A critique of statistical hypothesis testing in clinical research. *Journal of Ayurveda and Integrative Medicine* 2, 105-114.

- Raichle, M. E. (2006). The brain's dark energy. *Science* 314, 1249-1250.

- Roth, G. (1997). *Das Gehirn und seine Wirklichkeit. Kognitive Neurobiologie und ihre philosophischen Konsequenzen*. Frankfurt: Suhrkamp.

- Rawlins, M. (2008). *De Testimonio – On the Evidence for Decisions about the Use of Therapeutic Interventions. The Harveian Oration*. Delivered before the Fellows of the Royal College of Physicians of London on Thursday 16 October 2008. London: Royal College of Physicians.

- Tressoldi, P. E. (2011). Extraordinary claims require extraordinary evidence: the case of non-local perception, a classical and Bayesian review of evidence. *Frontiers in Psychology* 2(2), Art 117.

- Walach, H., Gaus, W., Haeusler, W., Lowes, T., Mussbach, D., Schamell, U., et al. (1997). Classical homoeopathic treatment of chronic headaches. A double-blind, randomized, placebo-controlled study. *Cephalalgia* 17, 119-126.

- Walach, H. (2000). Magic of signs: a non-local interpretation of homeopathy. *British Homeopathic Journal* 89, 127-140.

- Walach, H. (2002a). Response to Vithoulkas: Homeopathic fantasies about science, a meta-critique. *Homeopathy* 91, 35-39.

- Walach, H. (2002b). Reply to Vithoulkas and Oberbaum. *Homeopathy* 91, 189-191.

- Walach, H. (2008). Wissenschaft und Spiritualität. In G. Hüther, W. Roth & M. von Brück (Eds.), *Damit das Denken Sinn bekommt. Spiritualität, Vernunft und Selbsterkenntnis* (pp. 77-96). Freiburg: Herder

- Walach, H. (2009). Homeopathic pathogenetic trials – A summary of 20 years of reflection, data collection, and analysis. In C. Witt & H. Albrecht (Eds.), *New Directions in Homeopathy Research: Advice from an Interdisciplinary Conference* (pp. 43-66). Essen: KVC-Verlag.

- Walach, H. (2011). *Spiritualität: Warum wir die Aufklärung weiterführen müssen*. Klein Jasedow: Drachen Verlag.

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

Tags: research , meditation , medicine , teaching methods , pain management , spirituality , Walach , science

About the Author

Prof. Dr. Dr. Harald Walach
walach@europa-uni.de

Harald Walach is a clinical psychologist, philosopher, science historian, and a professor of complementary medicine research methodology and naturopathic medicine. He is also the director of the Institute for Transcultural Health Sciences (IntraG) at the European University Viadrina Frankfurt (Oder) Germany, director of their Master of Arts (MA) Program in Cultural Studies- Complementary Medicine, and director of the Samueli Institute European Office. His fellow experts of his field regard him as researcher with an excellent reputation and the magazine "Explore" cites him as "one of the most outstanding European researchers in the field of complementary and alternative medicine".