Truly personalised medicine: self-experimentation in medical discovery.

William Bains
Rufus Scientific
william [at] rufus-scientific.com

Published in *Medical Hypotheses* 70(4): 714-718 (2008)

ABSTRACT

Biomedical research need not be carried out solely by 'Them': distant, dissociated, enormously expensive institutions and companies. It can, and increasingly in the 21st century will, be carried out by 'Us', the informed non-professional. Conventional clinical trials treat humans with the same experimental model as laboratory rats - regarding them as mute, variable, unreliable material from which results must be obtained as fast as possible to maximize return on investment and patent life. The alternative is longer term, self-reported clinical studies of new treatments, based on the assumption that the experimenter is informed, intelligent and aware. A wide variety of new treatments for chronic disease are available, involving elements of diet, behavior, environment and non-prescription medication as well as ethical pharmaceuticals, and previous experience suggests that they can be enormously effective. The key is objective, quantifiable measures of outcome. These can be achieved with over-the-counter diagnostics for a variety of parameters, self-built test systems, and careful and systematic observations of symptoms. Hypothesis generation is a key part of this process.

Almost every article on the productivity of pharmaceutical research bemoans that the cost is rising while productivity is at best static (e.g. see [1]). The usual explanations are that this is because the hurdle a new therapy must jump to beat existing therapies is ever increasing. Drugs are not like pop records, where anything 'new' has a chance of success. Each new drug must be better than the existing pharmacopoeia, as if each new hit release had to beat the Beatles, Stones and Beachboys before it was even played.

But this should be achievable. Our knowledge of biology also advances, and simultaneously the technology to apply it expands in scope and affordability. The real reason is increasing hurdles of safety and efficacy that drugs have to pass, driven by a cycle illustrated in Figure 1. Clinical trials are getting more and more expensive, to provide data to pass ever-increasing regulatory requirements and an ever-increasing likelihood of litigation if a drug 'goes wrong'. In an attempt to make drugs that have a passing chance to overcome these hurdles, greater and greater investment is made to make discovery an quantifiably reliable, industrial process, and move it away from the Doctor-Patient (D-P) relationship. Both trial size and industrialization are hugely costly, and so can only be done by major research centers, usually those of global corporations, whose motivation is related to medical benefit only to they extent that it improves share price. Naturally this leads to distrust of the motivations of those involved, whether in academia or industry, and hence to demands for tightened regulation to protect patients from ambitious medics and profit-hungry companies, and a tendency to see deliberate malfeasance in every medical mishap. This leads to even more demands on clinical trials and regulation, and so the cycle turns.

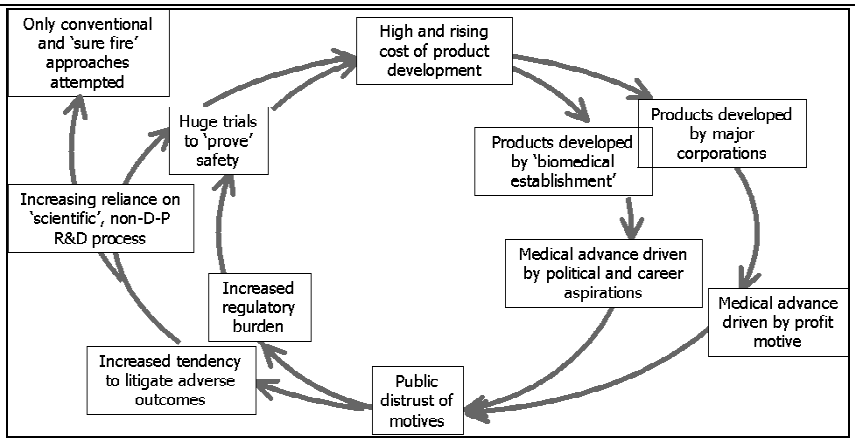

FIGURE 1
Cartoon of the process driving up drug R&D costs

To justify investment of $750M in a project you must show that it will succeed, in terms of the project not failing and in terms of the project providing leadership, expressed as patents, academic priority, market lead or other parameters. Paradoxically this leads to conservative projects - the radical is always likely to fail - and the use of the most complex, novel and above all expensive technology to get as far ahead of the competition as money can allow. This leads to conversion of messy, unreliable clinical medicine into clean, gleaming molecular biology. Several observers have commented on the collapse of 'clinical research' as a career[2]: dermatologists have become cell biologists, neurologists have become neuroscientists, and woe betide either of these is they put in a grant application without some gene sequencing in it. Recent moves to re-integrate clinical science into biological discovery focus on training clinicians in genetics, and bringing the shiniest of the high-tech equipment into special centers for Translational Medicine. Thus Translational Medicine conferences are full of discussions of PET fMRI, gene arrays and proteomics, far beyond the means of the GPs that see 95% of patients, and divorced from the simple clinical observations that resulted in the discovery of drugs as diverse as aspirin and viagra.

Because this is the way that biomedical research (especially drug research) is done, it is assumed that the features of this process are features that have to be part of the biomedical research process. These include:

i) that only professionals operating in established organizations can have the knowledge to identify new areas of medicines research
ii) that biomedical research can only be done using cutting edge technology, which is enormously expensive
iii) that only tests on huge numbers of people can validate a new approach.

None of these is true.

The first is obviously not true in the Google decade. Most scientific conferences in the US are now attended by a scattering of 'amateurs', and for the professional their level of expertise is astounding, usually exceeding the professional (i.e. paid) scientist in breadth and depth. Why? Because they care about the subject directly. Some career scientists pursue a specific line of research for love of the subject. Most pursue it because it is a subject for which grants are given and in which papers are published. The amateur cares nothing for papers or grants - she wants to know. Most have educated themselves over the net.

The second is only true if you chose the areas of research that are dependent on cutting edge technology. Seth Roberts (more on whom below) illustrates that there are many other areas, and I will touch on a few. Rees [3] also comments that the problem with current clinical research is not that there is a lack of funds (fMRI hardware does not come cheap, after all), but that the funds are directed at the wrong sort of research - research that requires lots of funds.

The third idea, however, is harder to dispose of. If a new intervention can be demonstrated only on large numbers of strangers, then the intervention has to be proven safe, manufactured to FDA-approved standards, monitored by disinterested observers and so on, all of which cost a fortune. This is not just an assumption by elderly and well-equipped laboratory directors. In my days as a venture capitalist there were few more effective ways to get a business plan thrown in the bin than to say 'I have tested this treatment on myself, and ...'.

But the reason behind such trials is the illusion that benefit can be demonstrated on people in the same way that it is tested on rats. The alternative approach is to test it on small numbers of people, maybe just one person, and rigorously determine that the effect on that person is real. The key is rigorous, objective testing. The huge clinical trial is based on the presumption that patients are essentially dumb - even if they do report on their own health, their reports are completely unreliable. This is how rats are tested. You cannot ask a rat whether it is feeling well, so you must observe a surrogate marker of how it feels, such as its movement, weight, histology, blood biochemistry or, in the classic LD50 tests, death. Such markers are actually quite unreliably linked to either 'wellbeing' or disease processes, so many rats have to be tested. You could test one rat, then try another treatment and test it, and then try another and so on, but that would take a long time and requires you to look after the rat very carefully between tests. The alternative is to test many rats in parallel and then throw them away, but that means even more rats, as rats are not all the same so the variation between them must be averaged. Precisely the same logic informs clinical trials: subjects are tested as if they were rats, and many tested in parallel because to test them in series would take too long and require too much control of their activities in the meantime - annualized returns on investment would fall, patents would expire and the investment that is needed for the trial would fail to earn a sufficient return. So huge trials of patients all tested at once are needed to pay for the trials of huge numbers of patients all at once.

Implicit in the statement of the problem above is an alternative solution. If the patient can report on their own health, and can try new treatments sequentially, then the need for external analysis of surrogates of wellbeing is removed, and much of the variation implicit in comparing Drug A on Patient 1 vs. Drug B on Patient 2 is removed. In short, the longitudinal, N=1 clinical trial should be a valid alternative to the current, hugely expensive approach.

Dr Seth Roberts is a high profile practitioner of this sort of medicine. His approach, and some of the results, are eloquently summarized in [4]. The key is objective measurement of health outcomes. If the outcome is weight this is fairly easy. If it is 'mood' it is harder - the success of Cognitive Behavioral Therapy shows that just thinking about being positive and happy can alleviate symptoms of depression. So the scientist-patient must be rigorous about their own analysis. Such rigor of analysis is rare in the 21st Century, when the passion with which a belief is held is considered to be justification of that belief in fields as diverse as the safety of GM crops, the effect of TV violence, and the justification for wars.

Within this framework, there are a huge number of measures that are cheap, accessible and relevant to health. Weight and height are obvious. There are 'Over-The-Counter' tests available in most Western countries for blood pressure, temperature, pulse, urinary glucose, blood cholesterol, peak breath flow etc., and these tests can be adapted to measure exercise tolerance, measurement of circadian rhythms, and a variety of metabolic parameters. Internet providers can sell a far wider range of tests to consumers. Elementary engineering can give tests for reaction time, expired breath volume, balance (for example, see Seth Roberts' blog for details of the $26 balance-measuring device), and via webcams such measures as pupil size, skin tone and so on. Software can measure a wide range of cognitive parameters. The list could go on and on. The

key to success is in using these simple measures in ways that convincingly control for observer bias and placebo effect.

There are also an enormous number of new therapies to be tested. It is common knowledge that medicine does not rely on 'drugs'. It relies on 'treatments' which are usually combinations of several drugs with other types of treatment - diet physical and mental exercise, surgery and others. Just looking at drugs alone, there are around 500 types of drug in the British National Formulary, most of which are sufficiently safe to be taken without medical supervision even if they can only be prescribed by a doctor. Another dozen or so classes are taken for recreation - legally or illegally. Most people taking drugs take more than one, and have more than one disease or syndrome: the elderly (who take most drugs) often take handfuls of pills a day for half a dozen conditions. Observation and follow-up of effects by the patients themselves, using rigorous measures of ill-health or well-ness, could mine this rich vein of data for new approaches that would never be considered by a medico-industrial system obsessed with huge trials, high-tech analysis and patentable outcomes. (It is ironic, therefore, that drug discoverers typically retire at age 60, just when they themselves start to become the most fertile ground for drug discovery.)

A key to this is not to trust an initial observation. Roberts' work on acne was triggered by an observation that the pills he was being prescribed for the condition did nothing at all, that the peroxide cream seemed to work, but with a delay. This lead to a hypothesis, that food could exacerbate acne and that oxidant creams helped: independent of the original analysis, he repeated the observation and validated it. Rees [3] comments that the majority of clinical advances in the last 20 years of dermatology have been made by individuals working outside the mainstream of academic research, but possessing a keen observational eye, strong, skeptical analytical skills and constant contact with patients. (Dermatology is a particularly ripe ground for such self-experimentation, as the skin is visible.)

Not all such tests will be successful. I thought that general inflammation would precede atopic symptoms after eating a food to which I was mildly allergic, and so that measuring body weight and temperature could predict onset of such symptoms (and hence allow pre-emptive use of lower doses of drugs), I have tracked a variety of physiological parameters as they fluctuated and correlated them with my atopy. Some interesting initial data did not pan out when I followed it up - my conclusion from months of data was that my measurements were not accurate enough. So, OK: define a better protocol or buy new equipment. Roberts has been self-experimenting for decades. This is the timescale needed. The idea, formented by the biomedical establishment, that 'breakthroughs' can be achieved with months' or, at most, a years' effort[3] leads us back to the rats. If the scientist-patient is to be their own control, then they must have patience to replace a huge budget as well as the rigor to replace statistics.

But a powerful way to avoid self-delusion is to have a theoretical underpinning for the new approach. Trying treatments at random runs the risk of generating apparently statistically robust results by chance alone - by definition 1 in 20 tests will produce a result that is 'significant' at the 95% level. If we measure 10 parameters once a day for 20 days, then even if they are entirely random we are likely to find that two of the parameters are correlated with $|r|>0.5$ (Figure 2). This looks meaningful, but is not. Running the observations for longer (say 100 days) reduces the correlation, but does not reduce its apparent significance. This is because in both cases we are in fact making 45 correlations (10C2), so of course 2 of them are likely to be 'significant at the 95% level' by chance. The more data we collect, the higher must be the threshold of significance. By contrast if we start from a single hypothesis, then 'p=0.05' significance level actually becomes significant. Roberts' weight loss regime arose from a hypothesis about how we had adapted to fluctuations in the food supply in the Pleistocene: from this hypothesis he made predictions, and then tested them by correlating weight with one or two input variables. Roberts' work also illustrates how a detailed knowledge of the literature informs and constrains hypotheses and speculation, again making actual observations more significant. In the Google decade, searching the literature is something that anyone can do.

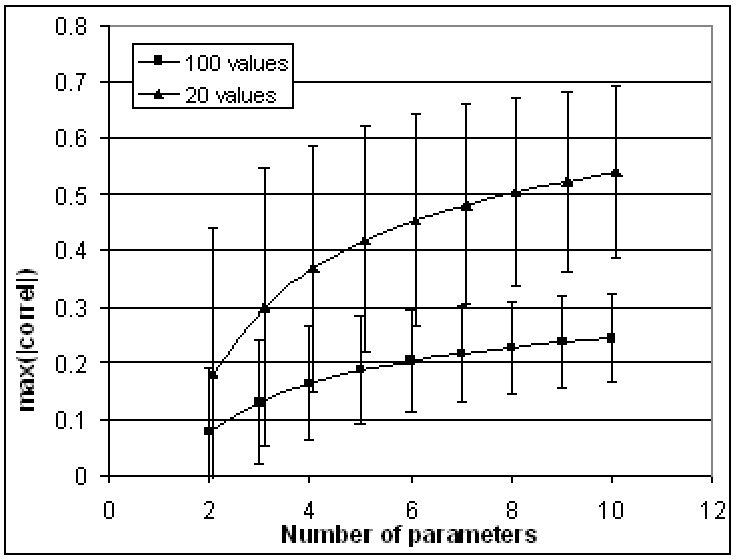

FIGURE 2
Correlation between the most correlated pair of several sets of random numbers. Y axis: the maximum absolute value of the correlation between a pair of sets of random numbers, selected as the most highly correlated from N sets of random numbers. X axis - N. Shown are the average 'best correlation' from 1000 sets of numbers, and 95% significance levels in that 'best correlation'. Squares - correlations between 100 values, triangles, correlations between 20 values.

This is where I believe that Medical Hypotheses can play a significant role. I have asked writers for this journal to conclude with how their hypotheses are testable before[5]. I would now urge them to go further. Testing can be in terms of gene chips and expression analysis and whole body MRI and all the other trappings of high-tech biomedical research, or in terms of large blinded clinical trials. But it can also be in terms of the individual. Writers here might consider not just how they can test their hypothesis in general, but how they as individuals could test their ideas, using a minimum of resources but the power of longitudinal, closely observed personal tests, and more generally how readers might test the ideas.

Lastly, there is an aspect of biomedical research which I have deliberately overlooked here. Testing diets, new uses and combinations of existing drugs, innovative nutraceuticals and 'personal management' regimes is relatively safe. Testing new therapeutic agents is a different matter. A substantial fraction of the cost of conventional drug discovery is making new drugs and confirming that they are safe enough to test on people, in terms of their underlying pharmacology and toxicology and in terms of their reliable (GMP) manufacture. It is plausible to suggest that an individual could try existing drugs on themselves. New medicines require a different paradigm, which is another topic for discussion.

ACKNOWLEDGEMENTS

My thanks to Jack Scannel and Jonathan Rees for discussions, and to Bruce Charlton for alerting me to Seth Roberts' work.

REFERENCES

1. Bains, W., Failure rates in drug discovery and development: are we getting any better? Drug Discovery World, 2004. 2004(Fall): p. 9 - 18.
2. Rees, J., Complex disease and the new clinical sciences. Science, 2002. 296: p. 698 - 701.
3. Rees, J., The Fundamentals of Clinical Discovery. Perspectives in Biology and Medicine, 2004. 47(4): p. 597 - 607.
4. Roberts, S., Self-experimentation as a source of new ideas: ten examples about sleep, mood, health and weight. Behavioral and brain sciences, 2004. 27: p. 227 - 288.
5. Bains, W., How to write up a hypothesis: the good, the bad and the ugly. Medical Hypotheses, 2005. 64: p. 665 - 668.