

Comment

Half right

Gregory A Petsko

Address: Rosenstiel Basic Medical Sciences Research Center, Brandeis University, Waltham, MA 02454-9110, USA. E-mail: petsko@brandeis.edu

Published: 30 September 2005

Genome Biology 2005, **6**:120 (doi:10.1186/gb-2005-6-10-120)

The electronic version of this article is the complete one and can be found online at <http://genomebiology.com/2005/6/10/120>

© 2005 BioMed Central Ltd

There's a wonderful moment in the equally wonderful 1973 film *'Sleeper'*, in which doctors in the year 2173 are discussing their new patient, Miles Monroe (played by the film's director, Woody Allen), who has just been awakened, like Rip van Winkle, from a 200-year hibernation (the result of a botched operation). "This morning," says one of the physicians of the future, "for breakfast, uh, he requested something called wheat germ, organic honey, and tiger's milk." To which another doctor remarks, "Oh yes. Those are the charmed substances that some years ago were thought to contain life preserving properties." "You mean," says the first doctor, "there was no deep fat? No steak, or cream pies, or hot fudge?" "Those were thought to be unhealthy," replies the other, "precisely the opposite of what we now know to be true."

Woody Allen would not be surprised at a thesis put forward by John Ioannidis, a Professor of Epidemiology who divides his time between University of Ioannina School of Medicine in Greece and Tufts University in the US, and neither, I suspect, would most Americans today. In an Essay just published in the Public Library of Science's journal *PLoS Medicine*, entitled "Why Most Published Research Findings Are False", (*PLoS Medicine* 2005, **2**: e124), Ioannidis asserts that "there is increasing concern that in modern research, false findings may be the majority or even the vast majority of published research claims. However, this should not be surprising. It can be proven that most claimed research findings are false." He goes on to give a few reasons why, arguing that claimed research findings may often simply reflect the prevailing bias, or be influenced by financial and other interests.

I won't summarize the statistical arguments he goes through to try to prove his point. They depend on models for bias and testing by several independent teams, and on my reading have a certain *ad hoc* character that makes me somewhat suspicious of them, but let's assume they may be valid. They lead him to some interesting corollaries, as follows. First, the smaller the studies, the less likely the research findings are to be true. Second, the smaller the effect sizes, the less likely the research findings are to be true. Third, the greater the

number and the lesser the selection of tested relationships, the less likely the research findings are to be true. Fourth, the greater the flexibility in designs, definitions, outcomes and analytical modes, the less likely the research findings are to be true. Fifth, the greater the financial and other interests and prejudices, the less likely the research findings are to be true. And sixth, the 'hotter' a scientific field, the less likely the findings are to be true.

I think items one through four should be subject to debate, but five and six sound logical to me, given human nature. Anyway, after six pages of basically arguing that every factor one could think of contributes to findings being false, Ioannidis never actually gives a final figure for what percentage of published results are wrong. In previous publications and interviews, however, the figure of somewhat over 50% gets bandied about, so let's be generous to ourselves and assume it's about half. That's what most science writers did when they wrote stories about this article - and did they ever write stories about it. Almost every important newspaper in the US carried reports with headlines screaming that half of all scientific research is false, many of them on the front page.

And it does seem as though every week there's a new report that contradicts a previous report. Fat is bad for you. No, it's not. Yes, it is. No, not all fat is, only certain fats. No, you need some of all fats. No, you don't. And so on. The yo-yoing in the popular press over the benefits versus the risks of birth control pills and hormone replacement therapy alone must have caused many women to run screaming to their doctors, who probably were just as confused as their patients. And given that the issues in most of the articles I'm talking about can be expressed as yes or no questions, the figure of 50% wrong sort of makes sense.

But is that really the case? Looking at the essay in more detail, I don't see how it can be. First of all, the title of the paper, "Why Most Published Research Findings Are False", gets my vote for the stupidest, most misleading title of the year. Nearly all of the examples are taken not from scientific

research in general but from medical research in particular, and most of them concern clinical trials of drugs or reports on the health benefits of various foods and diets. These do tend to be reported in 'yes or no' terms, so it's understandable why one might guess that half of them are false. And such studies suffer from a number of other factors that make them grist for the Ioannidis mill. They are frequently funded by organizations that have a vested interest in the outcome, so charges of bias are easier to make (though perhaps not to prove). They are often relatively small studies with a large number of variables. And they are being performed with the most difficult, pernicious, inhomogeneous experimental subjects in all of science: people. Finally, their results are usually reported in statistical terms, and many reporters, not to mention scientists, have only a rudimentary grasp, at best, of statistical concepts and pitfalls.

Few if any of these factors apply to many other areas of science. Hardly any of them apply to most branches of physics and chemistry or to basic research in general. Drug trials and tests of nutrients and environmental factors on human health are examples of highly targeted research with relatively absolute end-points. Basic research is not only more open-ended, it is a continuum. Studies tend to evolve rather than end, and intermediate results are publishable. Data are usually reproducible. Conclusions may be overturned as new data become available, but that doesn't make the research false, because the data are often right. I can't count the number of times I have gone back to the older literature and extracted enormously valuable data from a paper whose conclusions are no longer believed to be true. Classifying research papers as true or false belittles and grossly oversimplifies the way most fields work.

Besides, I wonder if it has occurred to the author of the essay that, if most published research findings are false, then his work is also more likely to be false than true, which would mean that most published research findings are true, which would mean that his are also likely to be true, which would mean that most published research findings are false, which would mean... what?

I can't help thinking we wouldn't be in this mess if so many scientists, especially in medical research, didn't feel it necessary to trumpet their findings in newspapers and popular magazines even when the real impact of the work is minimal. I have gotten so jaded with breathless statements of increased risk of dying from this or that which turn out to have only a 5-10% increase in the odds ratio that I have made it a policy not to get concerned unless the risk changes by at least a factor of two.

Of course, even that rule of thumb has to be applied carefully. It works most of the time because the odds of getting most diseases are pretty low if one is in good overall health, so a change of 10% in a probability of 1 chance in 500, say,

really doesn't amount to much. But there are situations where the risk is large enough that small changes in it matter. Neurodegenerative diseases like Alzheimer's and Parkinson's have risk factors that increase exponentially with age after one turns 60, and become quite high by the time one approaches 90, so things that modify those risks even by small amounts are worth attention. Polymorphisms in oncogenes and tumor-suppressor genes may also confer moderately increased risks of cancer in certain populations - an example is the I1307K single-nucleotide polymorphism in the *APC* gene, which is carried by about 1 in 20 Ashkenazi Jews and almost doubles their risk of colon cancer. I've become a big advocate for personalized medicine because things like that have important consequences: knowing one carries that mutation, for example, would seem to dictate earlier and more frequent colonoscopies than are usually recommended. As for claims for this vitamin or that type of diet, I've decided that most of those studies do nothing except increase by about 33% my chance of losing my lunch. Few fields are so beset with overinflated claims, misapplied statistics, and employment of scare tactics. Ioannidis is right about those papers, I bet. But his conclusions don't apply to other fields. And he shouldn't have implied that they do.

So let's see if I've got this right. Half of all medical research is right, and half is wrong, as long as this paper claiming that half of all medical research is wrong is right, but since there's an equal probability it's wrong, that would mean that the half of all medical research that was wrong might be right after all, unless of course all medical research is wrong, which would also be consistent with this paper being right. Right? Of course, I could be wrong.