

## Offline: What is medicine's 5 sigma?



Stuart Clarke/Rex Features

"A lot of what is published is incorrect." I'm not allowed to say who made this remark because we were asked to observe Chatham House rules. We were also asked not to take photographs of slides. Those who worked for government agencies pleaded that their comments especially remain unquoted, since the forthcoming UK election meant they were living in "purdah"—a chilling state where severe restrictions on freedom of speech are placed on anyone on the government's payroll. Why the paranoid concern for secrecy and non-attribution? Because this symposium—on the reproducibility and reliability of biomedical research, held at the Wellcome Trust in London last week—touched on one of the most sensitive issues in science today: the idea that something has gone fundamentally wrong with one of our greatest human creations.

\*



View Pictures/UGC via Getty Images

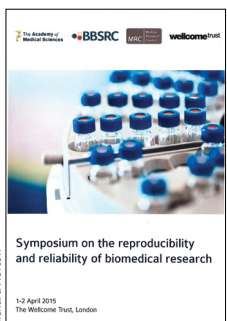
The case against science is straightforward: much of the scientific literature, perhaps half, may simply be untrue. Afflicted by studies with small sample sizes, tiny effects, invalid exploratory analyses, and flagrant conflicts of interest, together with an obsession for pursuing fashionable trends of dubious importance, science has taken a turn towards darkness. As one participant put it, "poor methods get results". The Academy of Medical Sciences, Medical Research Council, and Biotechnology and Biological Sciences Research Council have now put their reputational weight behind an investigation into these questionable research practices. The apparent endemicity of bad research behaviour is alarming. In their quest for telling a compelling story, scientists too often sculpt data to fit their preferred theory of the world. Or they retrofit hypotheses to fit their data. Journal editors deserve their fair share of criticism too. We aid and abet the worst behaviours. Our acquiescence to the impact factor fuels an unhealthy competition to win a place in a select few journals. Our love of "significance" pollutes the literature with many a statistical fairy-tale. We reject important confirmations. Journals are not the only miscreants. Universities are in a perpetual struggle for money and talent, endpoints that foster reductive metrics, such as high-impact publication. National assessment procedures, such as the Research Excellence Framework, incentivise bad practices. And individual scientists, including their

most senior leaders, do little to alter a research culture that occasionally veers close to misconduct.

\*

Can bad scientific practices be fixed? Part of the problem is that no-one is incentivised to be right. Instead, scientists are incentivised to be productive and innovative. Would a Hippocratic Oath for science help? Certainly don't add more layers of research red-tape. Instead of changing incentives, perhaps one could remove incentives altogether. Or insist on replicability statements in grant applications and research papers. Or emphasise collaboration, not competition. Or insist on preregistration of protocols. Or reward better pre and post publication peer review. Or improve research training and mentorship. Or implement the recommendations from our Series on increasing research value, published last year. One of the most convincing proposals came from outside the biomedical community. Tony Weidberg is a Professor of Particle Physics at Oxford. Following several high-profile errors, the particle physics community now invests great effort into intensive checking and re-checking of data prior to publication. By filtering results through independent working groups, physicists are encouraged to criticise. Good criticism is rewarded. The goal is a reliable result, and the incentives for scientists are aligned around this goal. Weidberg worried we set the bar for results in biomedicine far too low. In particle physics, significance is set at 5 sigma—a p value of  $3 \times 10^{-7}$  or 1 in 3.5 million (if the result is not true, this is the probability that the data would have been as extreme as they are). The conclusion of the symposium was that something must be done. Indeed, all seemed to agree that it was within our power to do that something. But as to precisely what to do or how to do it, there were no firm answers. Those who have the power to act seem to think somebody else should act first. And every positive action (eg, funding well-powered replications) has a counterargument (science will become less creative). The good news is that science is beginning to take some of its worst failings very seriously. The bad news is that nobody is ready to take the first step to clean up the system.

Richard Horton  
richard.horton@lancet.com



Richard Horton

Symposium on the reproducibility and reliability of biomedical research  
1-2 April 2015  
The Wellcome Trust, London