

Metascience: Reproducibility blues

Marcus Munafò

Nature **543**, 619–620 (30 March 2017) doi:10.1038/543619a Published online 29 March 2017

Marcus Munafò enjoys a stinging survey of unreliable findings in biomedical research.

Rigor Mortis: How Sloppy Science Creates Worthless Cures, Crushes Hope, and Wastes Billions

Richard F. Harris *Basic*: 2017. ISBN: 9780465097906

As scientists, we are supposed to be objective and disinterested, careful sifters of evidence. The reality is messier. Our training can give us only so much protection from natural tendencies to see patterns in randomness, respond unconsciously to incentives, and argue forcefully in defence of our own positions, even in the face of mounting contrary evidence. In the competitive crucible of modern science, various perverse incentives conspire to undermine the scientific method, leading to a literature littered with unreliable findings.

This is the conclusion of *Rigor Mortis*, a wide-ranging critique of the modern biomedical research ecosystem by science journalist Richard Harris. He describes how a growing number of claims over the past decade that many published research findings are false, or at least not as robust as they should be, has led to calls for change, and the birth of a new discipline of metascience.

He begins with the revelation in 2012 by Glenn Begley that only 6 (11%) of 53 'landmark' publications in preclinical cancer research could be confirmed by the biotechnology firm Amgen ([C. G. Begley and L. M. Ellis *Nature* **483**, 531–533; 2012](#)). Since then, numerous studies (most recently in psychology and cancer biology) have confirmed that failure to replicate published findings is the norm. The reasons are complex and contested. Harris identifies potential culprits, from the complexity of modern biomedical science to the limitations of tools and training, and perverse incentives in modern academia.

The scale of the problem is laid bare: apparently trivial methodological differences (such as how cells are stirred in culture, or the medium on which they're grown) can mean a complete failure to replicate results. Animal models often poorly predict results in humans; sample sizes can be too small to give reliable results; and some 12,000 studies may have used contaminated or misidentified cell lines. It is not just that the published research is unreliable — we may also be missing out on good drugs because poor preclinical data is an unreliable guide to whether to pursue human studies. The term “Eroom's law” (the reverse of Moore's law) has been coined to describe the worsening state of drug discovery. How much funding is wasted? Is the self-correcting nature of the scientific method functioning optimally? And, can we do better?



Stuart Wilson/SPL

Mice are often poor models for human therapies.

Harris introduces us to the growing field of metascience — the scientific study of science itself — and some of those working in it. These reproducibility firefighters are providing answers to such empirical questions, and identifying interventions. Robert Kaplan and Veronica Irvin at the US National Institutes of Health (NIH) showed that when the National Heart, Lung, and Blood Institute required preregistration of primary outcomes (the main outcome against which success should be judged) in clinical trials, the proportion of studies reporting a benefit fell from 57% to 8%.

Failure is a normal part of science, but dressing it up as success (for example, by presenting a secondary outcome as the primary outcome) is misleading. So is packaging exploratory, hypothesis-generating work as confirmatory, hypothesis-testing work. Unfortunately, with few ways to publish negative results, such practices are encouraged by incentives to present clean results with a compelling narrative, and be the first to do so.

Unsurprisingly, views differ on the reproducibility 'crisis'. Some believe we are in the dark ages; others, that attempts at direct replication are naive. The truth is probably in between, but the situation is sufficiently serious for key stakeholders to have begun to take notice and to introduce measures promoting robust design and transparent reporting. The NIH now has dedicated sections in grant proposals for applicants to describe how they will ensure their findings are robust, and *Nature* has introduced reporting checklists for submitted papers. There is growing interest in 'open science' — championed by the Center for Open Science in Charlottesville, Virginia — whereby elements of the research process (such as protocols, materials or data) are made publicly available. One positive outcome of the growth in metascience is that it has highlighted how every field typically does something very well, from preregistration to data sharing.

It is ironic that scientists in the pharmaceutical industry — often the target of opprobrium and worries about conflicts of interest — were among the first to raise concerns about the functionality of biomedical science. But it isn't surprising. They have incentives to be right — to make a correct 'go' decision on a compound that proves to be a successful treatment. Academic scientists, by contrast, are incentivized to publish first, to get grants and so on, but only rarely to get the right answer. In the words of Veronique Kiermer, executive editor at the Public Library of Science in San Francisco, California, "It actually pays to be sloppy and just cut corners and get there first". So what is good for scientists' careers may not be good for science. Simulations support this, suggesting that labs that do sloppy science will 'outperform' more- rigorous ones.

Harris makes a strong case that the biomedical research culture is seriously in need of repair. His focus is on preclinical research (and is rather US-centric), but he ends on a more optimistic note. The culture in various branches of biomedical science is changing, and incorporating lessons from other branches — preregistration of protocols, reporting checklists, and open data and materials. There is also cross-pollination of ideas between academia and industry. And funders and journals have begun initiatives to improve the quality of research.

Looked at in this way, biomedical research is not in crisis, but is embracing an opportunity to improve how it works, using scientific tools to understand the scientific process. Change takes time; *Rigor Mortis* shows that reproducibility issues are now mainstream, and that can only be good for science.