WORLD VIEW A personal take on events



Make journals report clinical trials properly

AUDIT AND

ARE THE BREAD AND

BUTTER OF GOOD

MEDICINE, AND

GOOD SCIENCE.

ACCOUNTABILIT

There is no excuse for the shoddy practice of allowing researchers to change outcomes and goals without saying so, says **Ben Goldacre**.

Science is in flux. The basics of a rigorous scientific method were worked out many years ago, but there is now growing concern about systematic structural flaws that undermine the integrity of published data: selective publication, inadequate descriptions of study methods that block efforts at replication, and data dredging through undisclosed use of multiple analytical strategies. Problems such as these undermine the integrity of published data and increase the risk of exaggerated or even false-positive findings, leading collectively to the 'replication crisis'.

Alongside academic papers that document the prevalence of these problems, we have seen a growth in 'technical activism': groups creating data structures and services to help find solutions. These include

the Reproducibility Project, which shares out the work of replicating hundreds of published papers in psychology, and Registered Reports, in which researchers can specify their methods and analytical strategy before they begin a study.

These initiatives can generate conflict, because they set out to hold individuals to account. Most researchers maintain a public pose that science is about healthy, reciprocal, critical appraisal. But when you replicate someone's methods and find discrepant results, there is inevitably a risk of friction.

Our team in the Centre for Evidence-Based Medicine at the University of Oxford, UK, is now facing the same challenge. We are targeting

the problem of selective outcome reporting in clinical trials.

At the outset, those conducting clinical trials are supposed to publicly declare what measurements they will take to assess the relative benefits of the treatments being compared. This is long-standing best practice, because an outcome such as 'cardiovascular health' could be measured in many ways. So researchers are expected to list the specific blood tests and symptom-rating scales that they will use, for example, alongside the dates on which measurements will be taken, and any cut-off values they will apply to turn continuous data into categorical variables.

This is all done to prevent researchers from 'data-dredging' their results. If researchers switch from these pre-specified outcomes, without explaining that they have done so, then they break the assumptions of their statistical tests. That carries a significant risk of exaggerating findings, or simply getting them wrong, and this in turn helps to explain why so many trial results eventually turn out to be incorrect.

You might think that this problem is so obvious that it would already be competently managed by researchers and journals. But that is not

the case. Repeatedly, academic papers have been published showing that outcome-switching is highly prevalent, and that such switches often lead to more favourable statistically significant results being reported instead. This is despite

NATURE.COM
Discuss this article
online at:

go.nature.com/8wdqhd

numerous codes of conduct set up to prevent such switching, most notably the widely respected CONSORT guidelines, which require reporting of all pre-specified outcomes and an explanation for any changes. Almost all major medical journals supposedly endorse these guidelines, and yet we know that undisclosed outcome-switching persists.

Our group has taken a new approach to trying to fix this problem. Since last October, we have been checking the outcomes reported in every trial published in five top medical journals against the prespecified outcomes from the registry entries or protocols. Most had discrepancies, many of them major. Then, crucially, we have submitted a correction letter, on every trial that misreported its outcomes, to the journal in question. (All of our raw data, methods and correspond-

ence with journals are available on our website at COMPare-trials.org.)

We expected that journals would take these discrepancies seriously, because trial results are used by physicians, researchers and patients to make informed decisions about treatments. Instead, we have seen a wide range of reactions. Some have demonstrated best practice: the *BMJ*, for instance, quickly published a correction on one misreported trial we found, within days of our letter being posted.

Other journals have not followed the *BMJ*'s lead. The editors at *Annals of Internal Medicine*, for example, have responded to our correction letters with an unsigned rebuttal that, in our view,

raises serious questions about their commitment to managing outcome-switching. For example, they repeatedly (but confusedly) argue that it is acceptable to identify "prespecified outcomes" from documents produced after a trial began; they make concerning comments that undermine the crucial resource of trial registers; and they say that their expertise allows them to permit — and even solicit — undeclared outcome-switching. Furthermore, they have declined to publish our response to their 850-word letter in the journal.

In our view, this is troubling. *Annals*' response helps to explain why studies repeatedly find outcome-switching to be hugely prevalent, despite policies to prevent it. But journal editors now need to engage in a serious public discussion on why this is still happening. We are providing specific worked examples to facilitate this discussion, and if our project is regarded as provocative, then that is misguided. Audit and accountability are the bread and butter of good medicine, and good science. Lives are at stake when subtle statistical signals of benefit and risk are sought in noisy, messy trial data. We hope that the structures of science really are in a state of flux, and still changing.

Ben Goldacre is a physician, author and senior clinical research fellow at the University of Oxford, UK. e-mail: ben.goldacre@phc.ox.ac.uk