



Why animal research needs to improve

Many of the studies that use animals to model human diseases are too small and too prone to bias to be trusted, says **Malcolm Macleod**.

This is the golden age of medical research. Around the world, scientists are spending more money, writing more papers and building more shiny institutes. Almost all grant applications suggest that a positive funding decision will support research that could lead to new treatments for condition X — usually a growing scourge of modern society.

Many medical discoveries have made real differences to the lives of a great number of people, but could the research be done better?

It seems self-evident that we should encourage high-quality work, but what makes for high quality is a matter of opinion, which hardens over the years into dogma on the assumption that the most established and most venerated got there for a reason, so if one wishes their good opinion then one should do as they did.

Take experiments that use animals to model human diseases. Empirical study of the quality of these experiments is an emerging field, but it does suggest that all is not well. The most reliable animal studies are those that: use randomization to eliminate systematic differences between treatment groups; induce the condition under investigation without knowledge of whether or not the animal will get the drug of interest; and assess the outcome in a blinded fashion. Studies that do not report these measures are much more likely to overstate the efficacy of interventions.

Unfortunately, at best one in three publications follows these basic protections against bias¹. This suggests that authors, reviewers and editors accord them little importance.

Other basic aspects of the design of experiments in animals also receive scant attention. In the face of pressures to reduce the number of animals used, investigators often do studies that are too small to detect a significant effect. To guard against such 'underpowered' studies, researchers should calculate the number of animals required to have a reasonable chance of detecting the anticipated effect given the expected variance of the data. Fewer than one in one hundred such publications report sample-size calculations².

Fewer still define beforehand the most important ('primary') outcome. As a result, they tend to report only the outcomes that happen to show statistical significance, reducing a rigorous, hypothesis-testing experiment to something more like observational research.

The tendency to publish only positive results is another flaw in animal research. Such bias not only prevents scientists from getting credit for high-quality research that happens to be neutral, but also gives a false impression of efficacy. My research has shown that in animal tests of treatments for focal cerebral ischaemia (a model for stroke), publication bias leads to an

overestimation of drug efficacy by about one-third³, increasing risk for both clinical-trial participants and the pharmaceutical industry.

Experimental approaches are not very different throughout the life sciences, so the biases are probably similar too. A scientist's environment is full of potential hazards, such as non-renewal of funding, and potential rewards — getting published and receiving grants. As long as cheap, underpowered studies are more likely to have exciting positive (if false) results than expensive, well conducted, large studies — and as long as journals don't seem to know the difference — the pressure will remain to do what everyone else does.

So we need to change the rules. If publication in high-impact journals continues to be a yardstick, then the review process must do much more to assess bias. The ARRIVE (Animal Research: Reporting *In Vivo* Experiments) guidelines⁴, endorsed by, among others, Nature Publishing Group, are a good start. But, as Don Quixote observed, the proof of the pudding will be in the eating.

There must also be better ways to publish neutral studies. If the focal cerebral ischaemia literature reflects the life sciences generally, then 16% of studies go unpublished, and tackling publication bias would increase the number of manuscripts published every year by 160,000. At current growth rates we would expect this increase anyway over the next four years, so sorting out publication bias should be possible.

At the very least, we should look for ways to register all experiments — so that investigators can receive credit for work done and so that those seeking to summarize what is known have access to all relevant data. Such a system could be flexible, with information embargoed for a time to protect intellectual property.

It is hugely distressing to hear highly motivated young scientists say that they would prefer to do their research 'properly', but that if they don't get more published from their PhD work they will never find a postdoc position. They feel forced to lower their standards. We owe it to them to create an environment in which the rewards for conducting high-quality research are more immediately apparent. ■

**CHEAP,
UNDER-
POWERED
STUDIES ARE MORE
LIKELY TO HAVE
EXCITING
(IF FALSE) RESULTS
THAN LARGE,
EXPENSIVE STUDIES.**

Malcolm Macleod is a clinical neuroscientist at the University of Edinburgh, UK, and a member of CAMARADES (Collaborative Approach to Meta Analysis and Review of Animal Data from Experimental Studies). e-mail: malcolm.macleod@ed.ac.uk

NATURE.COM
Discuss this article
online at:
go.nature.com/zdqwtq

1. Kilkeny, C. *et al.* *PLoS ONE* **4**, e7824 (2009).
2. Sena, E., van der Worp, H. B., Howells, D. & Macleod, M. *Trends Neurosci.* **30**, 433–439 (2007).
3. Sena, E. S. *et al.* *PLoS Biol.* **8**, e1000344 (2010).
4. Kilkeny, C. *et al.* *PLoS Biol.* **8**, e1000412 (2010).