

Animal models of neurological disease: are there any babies in the bathwater?

Malcolm Macleod,¹ H Bart van der Worp²

Findings from laboratory experiments are often offered up to neurologists and our patients as 'evidence' for exciting new treatments for some of the conditions we try to manage. This is especially true for stroke, multiple sclerosis and those relentlessly progressive neurodegenerative diseases where current treatments are limited, and which largely target symptoms rather than disease progression. Whether responding to an invitation to join a randomised controlled trial or to the patient who has heard on the *Today Programme* on the BBC that stem cell therapies may soon be available for their disease, how confident can we be that data from animal studies are relevant to our clinical practice?

There is consensus among most scientists and clinicians that much human benefit has come from understanding animal biology, and some consensus in the public mind that, that being so, animal experiments can be ethical. However, the delivery of human benefit to patients with neurological disease has been somewhat limited, and the gap between the laboratory and the clinic is large and well recognised. Even when pathological and functional characteristics of disease appear similar in animals and humans, it is difficult to model the multiple causations, insidious onset and sometimes decades-long progression in an animal whose normal lifespan is measured in months rather than decades.

These shortcomings have recently led to greater focus on the processes through which promising treatments from the laboratory are

developed for use in humans, and the current fashion for 'translational' research. However, many of our teachers would argue, with some justification, that they too conducted what we now badge as translational science. Moreover, the addition of the adjective 'translational' to a noun does not in itself guarantee a more sophisticated understanding of how we might more effectively develop treatments for human disease.

In the past 30 years or so, we have learnt a great deal about how the poor conduct and incomplete reporting of clinical trials can mislead, by compromising their internal validity (do the data really reflect what the treatment does in the trial population?), their external validity (are the findings relevant to other patients with the disease?) and by publication bias (do I have access to all of the data?). Sorting this out has been a slow and sometimes painful process, championed by organisations such as the Cochrane collaboration, but it has resulted in much more reliable information to guide patients and clinicians.

So where do laboratory experiments stand? Clinicians without bench experience might be forgiven for believing that in the precise world of the laboratory, heterogeneity is eliminated, measurement is precise, experimental design is perfect and errors—random or systematic—do not occur. While some scientists may indeed achieve this level of research excellence, most published research does not.

The CAMARADES group (the Collaborative Approach to Meta-Analysis and Review of Animal Data in Experimental Studies) has been

¹Reader, Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh, UK

²Assistant Professor, Department of Neurology, Rudolph Magnus Institute of Neuroscience, University Medical Centre Utrecht, Utrecht, The Netherlands

Correspondence to

Dr M Macleod, Centre for Clinical Brain Sciences, University of Edinburgh, Chancellor's Building, Edinburgh EH16 4SB, UK; malcolm.macleod@ed.ac.uk



trying to better understand how we can use data from animal studies to improve human health. Of most immediate concern is internal validity—the possibility that reports of animal experiments may not actually reflect what happens in animals, let alone humans, because of a failure to take simple measures to avoid bias.¹ Across almost 5000 animal experiments describing efficacy of candidate drugs in animal models of various conditions, including stroke, Alzheimer's disease, Parkinson's disease and multiple sclerosis, identified in systematic reviews and held on the CAMARADES database, a blinded outcome assessment was reported by only 1 in 5, randomisation by 1 in 6 and a sample size calculation by only slightly more than 1 in a 100. This is not good.

What's more, these failings are an important source of bias. For example, in animals, the free radical scavenger NXY-059 was more than twice as effective in non-randomised animal

stroke studies compared with randomised studies²; it was neutral in randomised clinical trials in stroke. In studies modelling multiple sclerosis, drug efficacy was one-third higher in studies which did not blind the assessment of outcome, and 'small' animal studies (3–4 animals per group) reported significantly greater effect sizes than 'large' studies (more than 10 animals per group).³ World-wise readers of *Practical Neurology* will not be surprised to learn that the bibliometrics commonly used to measure academic output do not capture these rather fundamental variations in study quality.

And, as with data from clinical trials, publication bias is important. Using information from over 1300 animal studies modelling stroke, we have estimated that at least 1 in 6 experiments remain unpublished, and that this leads to an overstatement of efficacy of at least 30%.⁴

So, as well as developing better models of neurological disease we also need to improve the way we use the models we already have, and develop an evidence based approach to understanding translational neuroscience. In other words, some 'reverse translation', taking study design issues now regarded as fundamental to good clinical research and applying them in the laboratory setting is essential.

A consensus statement of standards for stroke experiments⁵ has been followed by more general statements of reporting standards,¹⁶ and journals as well as funders are beginning to understand the pivotal importance of study quality. In the meantime, clinicians would be well advised to consider systematically the quality and validity of all the available laboratory evidence before rushing to clinical trial or to the radio studio.

Competing interests MM and HBvdW are founder members of CAMARADES, have been laboratory

scientists and as clinical trialists are consumers of the products of animal research.

Provenance and peer review Commissioned; not externally peer reviewed.

REFERENCES

1. **van der Worp HB**, Howells DW, Sena ES, *et al.* Can animal models of disease reliably inform human studies? *PLoS Med* 2010;**7**:e1000245.
2. **Macleod MR**, van der Worp HB, Sena ES, *et al.* Evidence for the efficacy of NXY-059 in experimental focal cerebral ischaemia is confounded by study quality. *Stroke* 2008;**39**:2824–9.
3. **Vesterinen HM**, Sena ES, French-Constant C, *et al.* Improving the translational hit of experimental treatments in multiple sclerosis. *Mult Scler* 2010;**16**:1044–55.
4. **Sena ES**, van der Worp HB, Bath PM, *et al.* Publication bias in reports of animal stroke studies leads to major overstatement of efficacy. *PLoS Biol* 2010;**8**:e1000344.
5. **Macleod MR**, Fisher M, O'Collins V, *et al.* Good laboratory practice: preventing introduction of bias at the bench. *Stroke* 2009;**40**:e50–2.
6. **Kilkenny C**, Browne WJ, Cuthill IC, *et al.* Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research. *PLoS Biol* 2010;**8**:e1000412.