EDITORIALS

Learning lessons from MVA85A, a failed booster vaccine for BCG

We must review how we use animal data to underpin clinical trials in humans

Malcolm Macleod *professor of neurology and translational neuroscience*

University of Edinburgh, Edinburgh, UK

The development and testing of new treatments is complicated. Many approaches are developed through the laboratory, and most involve animal studies—to give confidence in the biological concept underpinning the treatment, to show safety, or to show efficacy in animal models of the disease in question. These purposes are regulated differently: safety studies must be conducted according to established principles of good laboratory practice, but there is no such requirement for proof of concept or efficacy studies. The regulatory requirements around safety are, understandably, better developed for marketing authorisation of new treatments than they are around approvals for clinical trials.

The development of MVA85A for the prevention of tuberculosis (TB) gives us an opportunity to reflect on these processes. This promising intervention was designed to boost the effectiveness of the well established BCG vaccine. It was reported to show efficacy in laboratory studies in animals but did not show benefit when tested in a large clinical trial in infants.[1] In the linked feature (doi:10.1136/bmj.j5845),[2] Deborah Cohen raises important questions about the development of MVA85A and, in particular, the extent to which emerging findings from animal studies were communicated to clinical trial regulators, funders, ethics committees, and potential participants.

It seems that while the human clinical trial was in the late stages of preparation, an experiment suggested that macaque monkeys given MVA85A in addition to a BCG vaccine had a higher mortality than those given BCG vaccine alone.[3] In this experiment, mortality endpoints were reached when monkeys became unwell enough to require euthanasia. Like most laboratory experiments, this one was too small to say whether the increased mortality was real or occurred by chance alone, but it flagged a potential problem and it would have been prudent to have at least made the clinical trial regulators aware of this finding at the earliest opportunity. Indeed, the European Medicines Agency's *Guidelines on Good Clinical Practice* recommend that the investigator brochure includes "all relevant non-clinical … studies."[4]

That these data were apparently not included in the initial investigators brochure (May 2008—six months after the

experiment ended), while supportive animal data were included, is a cause for concern. The data on mortality were provided later, in a November 2008 update to the brochure, seven months before the human clinical trial began, as part of a larger dataset from the macaque study.

Some uncertainty surrounds the purpose of the primate experiment, described in the initial research contract as an efficacy study but later characterised as the development of a new experimental model. Quite when the purpose changed is not clear, and a protocol registered before the animal study began[5] is not available. Without one, we do not know if the purpose changed before or after the data were available for analysis, and that is critical to assessing the importance of the findings. The study as published is characterised as a test of the model rather than as an efficacy study. Of course, investigators have a duty to change experimental plans when circumstances change if this increases the usefulness of the work; but any such change must be transparent and justified, so that research users can interpret findings accordingly.

Weak evidence

In fact, a recent systematic review and meta-analysis of animal studies has since shown that the MVA85A booster was associated with a small (3%) and non-significant reduction in bacterial load in animals infected with TB and a non-significant increase in risk of death when compared with BCG alone (risk ratio 1.45, 95% confidence interval 0.23 to 8.98).[6] Only one of the five included studies that reported risk of death—a guinea pig study with intradermal vaccination—found a statistically significant benefit from MVA85A plus BCG compared with BCG alone.

None of the eight published experiments included in that review reported a sample size calculation or blinding of the assessment of bacterial load or the decision to euthanise, and only one reported random allocation of animals to experimental groups. Unfortunately, this kind of weak evidence base is common in animal studies underpinning human trials. When animal data are pivotal in developing new treatments, we should insist on more robust evidence before proceeding.

malcolm.macleod@ed.ac.uk

The evaluation of MVA85A included data from only 192 animals, and just 40 primates. The idea that this is enough efficacy data in animals to support further large scale testing in humans is widely held but somewhat implausible and in need of review.

Much of the animal data on MVA85A came from a single research group, but the effects observed in animal studies can differ substantially among laboratories.[7 8] Much of the variation reflects true biological heterogeneity, so combining findings across several sites might increase confidence that effects observed in animals are robust enough to justify their clinical exploitation.[9]

Of course, it would be entirely reasonable to proceed to clinical trials on the basis of indirect evidence of potential efficacy (such as the immunogenic response to vaccination in animal models); but when the criteria by which this judgment will be made are not articulated clearly in advance there is a real danger that investigators will be seduced by the positive and rationalise away the neutral or negative, leading to premature clinical trials that may expose participants to harm. We need to develop better and more systematic ways to establish when a drug is ready for clinical trials in humans—and importantly, when it is not.[10]

Need for cultural change

The story of MVA85A also raises questions about how researchers and institutions respond to criticism. One approach sees criticism as a challenge that must be rebutted and resisted. An alternative characterises criticism as an opportunity to better understand where things might have gone wrong, to allow everyone to learn from what has happened, and to improve how we develop much needed new drugs.

While the second approach is at the heart of the ambition for self correcting and collegiate scientific endeavour, the first draws more from competitive environments, where "winning" isn't the main thing, it's the only thing. The current scientific ecosystem sees publication in high impact journals and the award of high value grants as ends in themselves. There is much discussion about how this might change, but until our institutions recognise that their core purpose is to produce research of value to society they risk a slow decline in their reputation, and possibly a faster and more serious erosion of public trust in science. In these troubled times, that public trust is more important than ever.

1   Tameris MD, Hatherill M, Landry BS, et al. MVA85A 020 Trial Study Team. Safety and efficacy of MVA85A, a new tuberculosis vaccine, in infants previously vaccinated with BCG: a randomised, placebo-controlled phase 2b trial. *Lancet* 2013;381:1021-8. doi:10.1016/S0140-6736(13)60177-4 pmid:23391465.
2   Cohen D. Oxford vaccine study highlights pick and mix approach to preclinical research. *BMJ* 2018;360:j5845. doi:10.1136/bmj.j5845.
3   Sharpe SA, McShane H, Dennis MJ, et al. Establishment of an aerosol challenge model of tuberculosis in rhesus macaques and an evaluation of endpoints for vaccine testing. *Clin Vaccine Immunol* 2010;17:1170-82. doi:10.1128/CVI.00079-10 pmid:20534795.
4   EMA. Guideline for good clinical practice E6(R2). 2015. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2015/08/WC500191488.pdf.
5   Macleod MR, Howells DW. Protocols for laboratory research. *Evidence-based Preclin Med* 2016;3: e00021.doi:10.1002/ebm2.21.
6   Kashangura R, Sena ES, Young T, Garner P. Effects of MVA85A vaccine on tuberculosis challenge in animals: systematic review. *Int J Epidemiol* 2015;44:1970-81. doi:10.1093/ije/dyv142 pmid:26351306.
7   Crabbe JC, Wahlsten D, Dudek BC. Genetics of mouse behavior: interactions with laboratory environment. *Science* 1999;284:1670-2. doi:10.1126/science.284.5420.1670 pmid:10356397.
8   Voelkl B, Würbel H. Reproducibility crisis: are we ignoring reaction norms? *Trends Pharmacol Sci* 2016;37:509-10. doi:10.1016/j.tips.2016.05.003 pmid:27211784.
9   Bath PM, Macleod MR, Green AR. Emulating multicentre clinical stroke trials: a new paradigm for studying novel interventions in experimental models of stroke. *Int J Stroke* 2009;4:471-9. doi:10.1111/j.1747-4949.2009.00386.x pmid:19930059.
10   Mogil JS, Macleod MR. No publication without confirmation. *Nature* 2017;542:409-11. doi:10.1038/542409a pmid:28230138.