What do animal experiments really tell us?

Robert Matthews*

Do animal models reliably predict toxic effects in humans, or are they actually blocking development of vital new drugs? Two recent major health scares have reignited the old debate.

Two huge industries affecting the lives of millions are currently in the midst of major health alerts. Concern over serious side-effects have cast a long shadow over promising new painkillers developed by the pharmaceutical industry known as COX-2 inhibitors. Evidence linking the drugs to an increased risk of heart attacks led the US giant Merck to voluntarily withdraw its version, known as Vioxx, from the market last September, and an investigation by the US Food and Drug Administration (FDA) raised concern about similar drugs.

Then in February it was the turn of the UK food industry, with the discovery of traces of a banned dye known as Sudan I in a sauce made by Premier Foods, a leading UK supplier. In the ensuing health alert, the UK Food Standards Agency (FSA) found that hundreds of products had been inadvertently contaminated by the dye, which has been linked to cancer.

As the initial furore starts to fade, both these health scares are being seen primarily as wake-up calls to both industry and regulators about the monitoring of product safety.

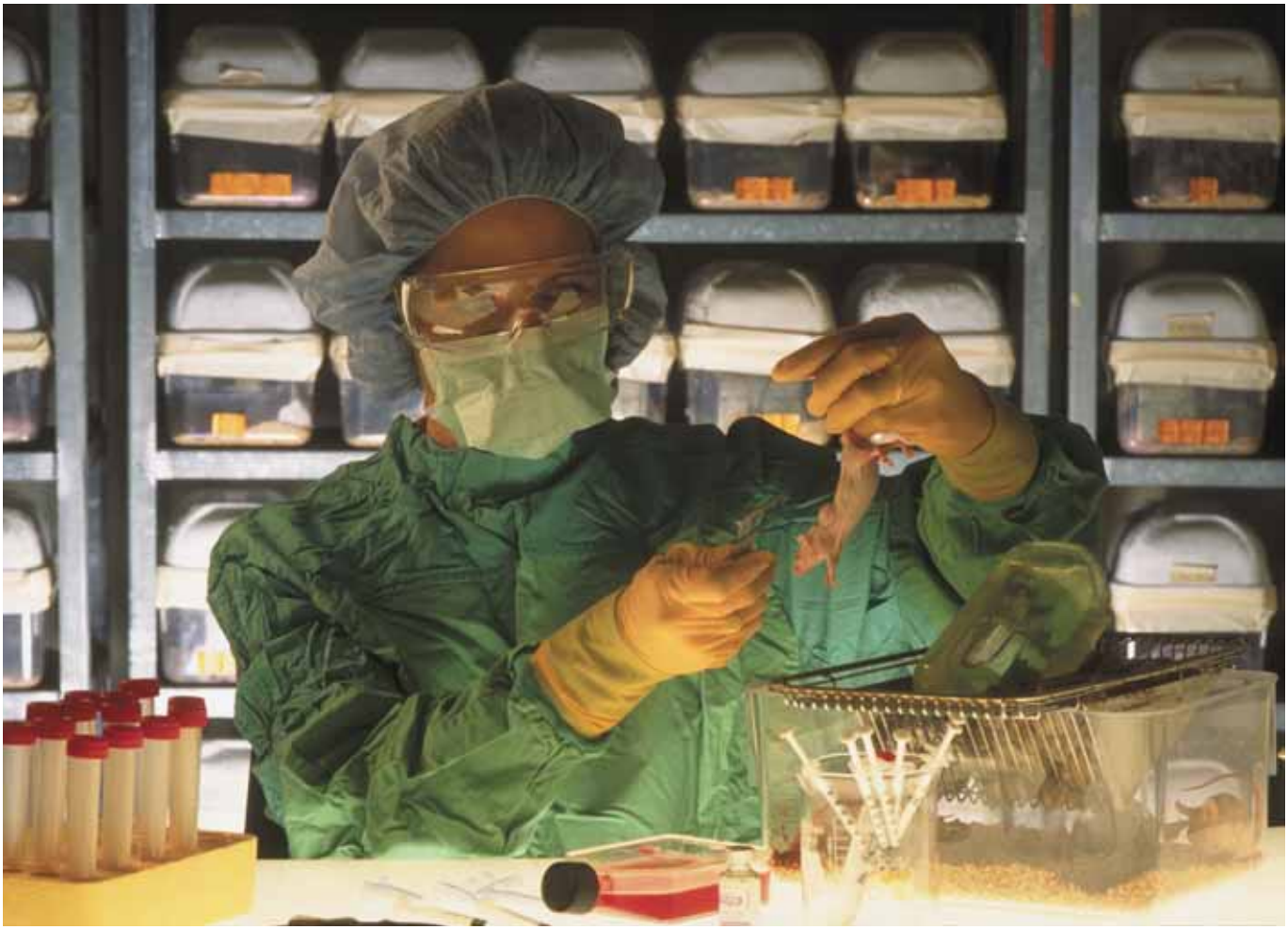
In the case of COX-2 inhibitors, the FDA is allowing some to remain on the market – albeit with much sterner safety warnings to protect those most at risk from side-effects. Meanwhile, while shops and supermarkets in the UK hunt down produce contaminated with Sudan I, the FSA has continued to stress that the risks involved are "very small".

As well it might, for it is now clear that the scientific case against Sudan I is far from compelling. Laboratory safety tests involved feeding rodents with levels of Sudan I equivalent to human consumption of the sauce that triggered the scare at a rate of three tonnes a day for two years. Even after such gargantuan exposure, the animals failed to produce consistent evidence of a cancer risk. Other tests hinted at links with bladder and liver tumours – but only after the dye was injected directly into the organs of laboratory animals. While the scientific basis for both the Sudan I and COX-2 inhibitor health scares may be contentious, they have highlighted the need for close surveillance and prompt action if problems emerge. At the same time, however, an even more fundamental question has gone begging: just how reliable are animal tests of product safety?

In the case of food safety, the relevance to humans of animal tests involving colossal intakes or direct injection into organs is clearly questionable. The use of animals in drug safety testing raises altogether more complex issues, however – as the COX-2 painkillers controversy shows.

In line with standard practice, Vioxx and the other drugs were tested in at least two different types of animal before entering clinical trials with humans. One of the key aims of such "pre-clinical" testing is to detect signs of serious side-effects. In the case of the COX-2 drugs, the animal testing failed to warn of the cardiovascular effects that have prompted the current furore. Indeed, several animal studies suggested the drugs would actually reduce the risk of such side-effects.

So what went wrong? Antivivisectionists have been quick to voice their standard objection: animals are not humans. For all its familiarity, it is an argument that does have relevance to COX-2 inhibitors like Vioxx. In 2000, barely a year after the launch of Vioxx, a study of over 8,000 patients suggested that those taking the drug faced a significantly



STEVE CHENN / CORBIS / CONTRASTO

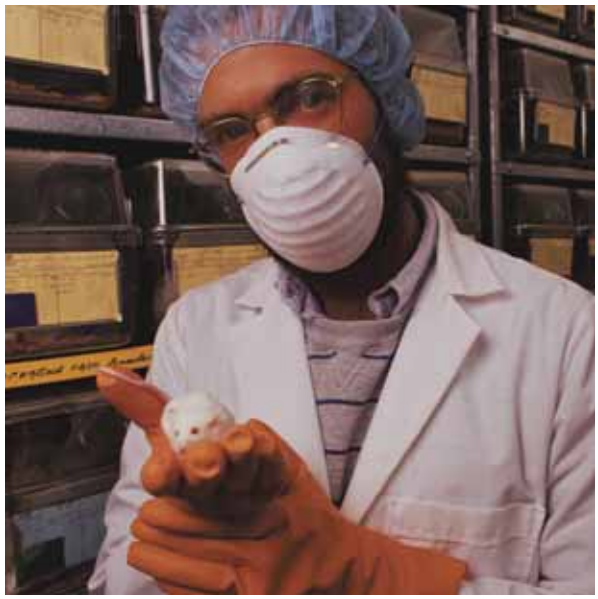
Animal tests may be blocking the development of many safe and effective treatments

increased risk of heart attack. Yet subsequent animal-based research continued to suggest such drugs could reduce the risk – prompting even Merck's own experts to concede in the *American Heart Journal* that "The relevance of these animal models in predicting effects in humans is uncertain."

It is becoming clear that such uncertainty extends far beyond one class of blockbuster drug. The leading journal *Nature Reviews: Drug Discovery* last year published a review of the evidence that animals are reliable predictors of toxic effects in humans. The authors found that the evidence was "fragmentary", with the few published studies pointing to "significant over- and under-prediction of adverse effects from animal studies that varies with the particular organ or system."

The review also highlighted the lack of basic data needed for a scientific assessment of animal testing, such as measures of predictive power and their statistical significance.

As it stands, the evidence suggest animal tests may be unduly sensitive,



ED KASHI / CORBIS / CONTRASTO

The animal testing failed to warn of the cardiovascular effects that have prompted the current furore

wrongly predicting toxicity in compounds that are in fact harmless to humans. If so, it would be an ironic twist to the widely-held belief that tests of animal are crucial to the advancement of medicine, as they may in fact be blocking the development of many safe and effective new treatments.

Yet in the absence of large-scale studies comparing drug responses in animals and humans, it is impossible to know. Supporters and critics of animal testing continue to trade anecdotes of individual successes and failures, more systematic studies typically being so small they lack statistical credibility. In another irony, the drive to minimise the use of animals has compelled researchers to draw huge conclusions from meagre evidence. For example, the studies designed to probe the effect of COX-2 inhibitors on cardiovascular risk typically involved fewer than 20 mice. The authors of last year's review called on both regulatory bodies and drugs companies to publish data currently languishing in their files. Whether the outcome will confirm or confound the view that animals usefully predict human reactions remains to be seen. What is clear is that, given the current paucity of systematic evidence, it is not necessary to be a placard-waving protestor to harbour doubts about the validity of animal testing.

DETECTING NASTY EFFECTS IS (MUCH) HARDER THAN IT SEEMS

The health scares over COX-2 drugs and the food dye Sudan I have highlighted the challenge of assessing health risks from limited data. While studies involving huge numbers of patients or laboratory animals are clearly better at detecting side-effects than small ones, they are also far more expensive and time-consuming. Worse, the ability of a study to detect risk does not increase pro-rata with size: to double the sensitivity, the required number of patients quadruples.

Worst of all, estimating the required numbers demands some guesstimate for the likely level of risk – and a bad guess raises the danger of the study being "underpowered", lacking the numbers needed to detect the true level of risk.

One solution is to set up a trial so large that it is sure to have a reasonable chance of detecting serious side-effects in one patient out of every *N* taking the drug. Statistical theory then shows that a comparison of 4 times *N*-squared patients taking the drug with the same number taking a placebo will do the trick. The bad news is that for blockbuster drugs like Vioxx, side-effects affecting just 1 in 1,000 patients constitute a major health alert – and detecting that level of risk demands a study involving millions of people. The only way of acquiring such vast numbers is for pharmaceutical companies and regulators to keep drugs under close surveillance long after approval.

* Robert Matthews is Visiting Reader in Science at Aston University, Birmingham, UK

This is an edited version of an article first published in the *Financial Times* on 4 March 2005