How Designing Better Experiments Can Reduce the Number of Laboratory Animals Used in Biomedical Research

Dr Michael F W Festing, FRAME

Animal research in the UK is tightly regulated under the Animals (Scientific Procedures) Act 1986. The system involves licences, detailed discussion of every project, and unannounced visits by Home Office inspectors as outlined by Lord Sainsbury in the Autumn 2003 edition. Yet there is still room to reduce the number of animals used in biomedical research by better design of experiments and using statistical methods more effectively. It would also improve the quality of the research saving both money and scientific resources.

A well designed experiment usually involves a comparison of several groups of animals given different experimental treatments. The aim is to identify the effect of the treatments on the animals. The good experiment should: ensure that the *only* difference between groups is due to the treatmentl; be *powerful* enough to detect any biologically important effects; be simple enough to minimise the risk of making mistakes; lend itself to statistical analysis and be economical with animals and scientific resources.

The power of an experiment depends largely on having uniform animals (ie of similar age, weight and genetic composition) and on the number used. This number has, until recently, been a matter of tradition and guess-work. Groups of about eight animals per treatment are common but when many treatments are involved this seriously over-estimates the number needed.

Group size

There are better methods for determining group size. According to the Resource Equation method the total number of animals, for measurement outcomes, should be the number of treatment groups plus 10 to 20 additional animals, rounded to equal numbers per group. Where there are more than 20 groups, each group should contain two animals. The Power Analysis method depends on the variability of the animals, the magnitude of the treatment response and the chance of reaching a wrong conclusion. Both methods reduce the guess-work.

Too many animals used

Experiments often use too many animals. A small survey of 27 UK scientific papers found that the number of animals used per experiment ranged from five Rhesus monkeys to 288 mice. The latter experiment involved 144 mice of each of two strains. The aim was to see if the strains differed in blood levels of three pharmacological preparations at six times of the day. It involved 36 groups of eight mice per group. However, it is unwise to do such a large experiment without having an idea of the outcome. Blood levels were not measurable for two of the pharmacological treatments. A pilot study using, say, the three pharmacological treatments, two times, two strains and two mice per group (24 mice) would doubtless have shown that the two pharmacological treatments gave undetectable blood levels. A second experiment could then have been done to see if the mouse strains differed using the single measurable pharmacological treatment, say at three times of day, using four mice per group or a total of 24 mice. This strategy would have used 48 mice, saving 240 mice or freeing resources for more experiments, thereby speeding research. Moreover it would have been more likely to reach a correct conclusion.

Faulty design in other respects and incorrect statistical methods rendered the conclusions reached by these authors unsafe.

Other experiments involving 88 rats, 102 rats and 64 mice could each have been done with about half these numbers. In each case the authors used simple but inappropriate statistical methods to try to analyse the data from complex experiments with many treatment groups. Several other papers had design errors or failed adequately to explain their methods, so it was impossible to judge whether they had been done correctly. In eleven cases it was not even clear how many animals had been used.

The inbred strain

Fourteen papers used rats but only one used an inbred strain. These strains are like clones of genetically identical individuals and have been available for many years. Their uniformity leads to more powerful experiments and they have several other useful characteristics. Their use is often essential and crucial yet, although at least eighteen Nobel prizes have been awarded for research necessitating the use of these strains, many research workers seem to be unaware of their valuable properties.

Scientists Failing

All papers discussed here were peer reviewed, implying that far too many scientists are failing in their understanding of experimental design and statistics. Fortunately the major funding organisations are now aware of the problem and the Medical Research Council, working through Centre for Best Practice for Animals in Research, has recently set up a working party of stakeholders to consider what needs to be done.