Reduction by Well-defined Objectives

Derek Fry

Animals (Scientific Procedures) Inspectorate, Home Office, P.O. Box 31, Shrewsbury, Shropshire SY3 7WN, UK
E-mail: *derek.fry@homeoffice.gsi.gov.uk*

Summary — Examination of programmes of experimental work involving live animals indicates scope for worthwhile reduction in animal usage by identifying more clearly the objectives of the different stages of a programme. Having a defined first objective for new or transferred methodology as "to optimise the methods", rather than an exploratory "try it with a few animals and modify" approach, is more likely to obtain the conditions for largest divergence between the effect to be measured and the background variability. Where several experiments will use the same methodology, even small improvements in this "signal to noise" ratio may be worthwhile, as the reduction is cumulative. Subdivision to more precise objectives helps concentrate the animal usage. For example, separating "to determine the time of a distinct peak" from "to determine the size of the peak" avoids using the numbers needed to achieve the latter for every time point at which the peak might occur. Similarly, setting a series of objectives for a programme that lead to clear decision points allows setting of clear criteria for achievement at each stage, as well as for the design and analysis to be the most efficient. This paper gives illustrations of the applicability of these arguments and estimates of potential reduction.

Key words: *experimental design, objectives, research strategy.*

Introduction

Judging whether a reduction in numbers of animals used in proposed experimental studies might be achieved is part of the work of Animals (Scientific Procedures) Inspectors in the UK. Every inspector advises on applications for licences for 5-year programmes of experimental work on animals, in addition to inspecting licensed work in progress. There is a legal requirement for an inspector to be consulted before animal work is licensed, and he or she has to advise on whether each programme involves minimal animal use.

As each inspector provides advice on some 4% of the UK experimental animal work, this means making an assessment on how animal usage can be minimised for a wide range of types of experimental studies. A number of the proposed studies would not use minimal numbers, and some examples will be given of proposals where more thought on the objectives should allow fewer animals to be used without increasing the severity or complexity of the work. These are illustrative of different types of approaches and based on actual proposals, but with details omitted or changed to preserve confidentiality.

In the UK, the *Animals (Scientific Procedures) Act 1986* requires that the authorities "shall not grant a . . . licence" for an experimental programme "unless . . . satisfied . . . that the . . . procedures to be used are those which use the minimum number of animals, involve animals with the lowest degree of neurophysiological sensitivity, cause the least pain, suffering, distress or lasting harm, and are most likely to produce satisfactory results" (1). An inspector advises on this before the licence is granted.

In addition to this scrutiny before granting the licence, during the 5-year programme that the licence authorises, a condition placed on the licence at the time it is issued constrains animal numbers used (2). This condition states:

For any procedure, the degree of severity imposed shall be the minimum consistent with the attainment of the objectives of the procedure, and this shall not exceed the severity limit attached to the procedure. The minimum number of animals of the lowest neurophysiological sensitivity shall be used in procedures causing the least pain, suffering, distress or lasting harm.

This condition places minimising numbers in the context of the objectives and makes clear that a balance has to be struck between reduced numbers and severity involved. On visits of inspection an inspector will look at, among other things, whether the work done under licence conforms to this condition.

Research Strategy and Experimental Design

Experience in the advising and scrutiny that is much of a UK inspector's job indicates that there is scope for both refinement and reduction in an experimental programme at two levels:

1. in the design of the programme — strategy; and

2. in the design of the individual experiments.

Refining strategy for an experimental programme depends on identifying clear objectives/decision points. With these set, it may be apparent that unclassified or mild work or lower species studies can provide part of the data sought, or that pilot studies are needed to show the nature and/or timing of adverse effects and to give a more reliable estimate of variability.

A workshop at the Third World Congress (3) produced the following ideas for steps in devising experimental strategy:

1. Breakdown of the problem into a series of key questions to investigate.

2. Determination of the objectives at each stage and decision criteria as to whether to proceed.

3. Determination of whether there are non-animal alternatives for each stage.

4. Identification of scientific measures to be taken — data to be acquired.

5. Consideration of how to maximise signal to noise ratio for each measure.

6. Identification of suitable pilot experiments to determine unknowns (especially variability) and humane endpoints appropriate to the objectives.

7. Identification of the best experimental design for each stage in terms of replacement, reduction and refinement.

8. Decisions on suitable statistical analysis.

9. Determination of numbers needed in the various groups.

The first step is identifying the key questions and setting appropriate objectives. In the following examples, the objectives of the proposed work could be redefined in a way that should lead to the same final result with fewer animals used.

Better-defined objectives example 1: separating different experimental stages

In the first case, the stated objective was to determine the effect of exercise on sympathetic nerve discharge. The sympathetic nerves involved were detailed in the proposal, and there was a clear scientific value in studying the link to exercise. To achieve this objective, major surgery was proposed to insert the electrodes, and recording was to take place some days later when the conscious animal was exercising.

This objective can be subdivided into three:

1. To optimise methods.

2. To determine the effect of muscle activity.

3. To determine the effect of conscious exercising.

The first of these places the emphasis on getting the best conditions for obtaining reliable results in the definitive studies — which operative approach to use, where best to place the electrodes, how to get good recording through the attached leads, and so on. It can be achieved using animals entirely under general anaesthesia from which they do not recover. The second focuses on whether effects on sympathetic discharge may be seen during passive movement or reflex muscle activity, and whether responses might occur solely from joint movements, or artefactually. This second objective could also be achieved using animals under terminal general anaesthesia. By the time the work passes to the third objective, meeting which would require recovery from the major surgery, the researchers would be practised in the methods, and possibilities for artefacts would have been explored, as well as some useful data gathered.

This staged approach should mean that full use could be made of animals subjected to the greater severity of the recovery period, as most of the learning phase will be over. The overall severity of the studies should be less, and how clear objectives can produce useful refinement like this has been explored previously (4). However it is also likely that the total numbers of animals used will be reduced as well. The optimisation should maximise the signal to noise ratio, enabling a smaller difference in signal to be detected, and the practice in the technique gained in the preliminary work should reduce variation due to operator inconsistency in the final phase. The passive and reflex exercising should give an indication of the variability of the signal.

As the estimate of numbers needed depends on the square of both the difference in effect (signal) that can be reliably detected and the variability expected, small changes in these make a big difference to group size estimates for definitive studies. For example, a group size estimate of 15 based on a coefficient of variation of 12% would reduce to 10 if the coefficient was 10%, and a 40% improvement in detectable difference between treated and control could half the group size estimate. Provided there are several studies on conscious animals, looking at perhaps different durations or intensities of the exercise undertaken, the numbers used in the preliminary work should be more than compensated for by the greater reliability and better estimate of group size in the later phase.

Better-defined objectives example 2: separating different aspects of an investigation

In the second example, the objective was to determine the peak of a growth factor response following injury, and the experimenter proposed to kill groups daily from days 1 to 20 following injury. Repeated measures were not possible, as the scientific data could only be collected post-mortem, but a single animal provided enough material for an assay. The methods were already optimised, and a good estimate of variance could be made. A group size of 20 was proposed, following a statistician's advice. For the experimental groups, 400 animals would be needed with 200 controls proposed (every second day).

However the main objective can be seen as two separate ones, namely:

1. To determine the timing of the peak.

2. To determine the size of the peak.

To determine the timing alone, it is not necessary to have the number of animals that are needed to give a reliable estimate of the size of the effect at a particular time point. It is sufficient to pick out the time points when the measurement is above the background range. Provided the increase is spread over more than one successive time point, it should be satisfactory to kill one animal at each time point. In this case, that would mean killing an animal a day over the 20 days of interest, that is, 20 animals in all. If the timing was not clear, as might occur with a peak occurring over one day only, the sequence could be repeated, using a total of 40 animals. If necessary a further repetition with 20 animals could be run. If nothing were seen after three runs, it would be likely there was no peak worth detecting. Should a faster result be required, or should the results from a run take some while to process, then an experiment with three animals per group over the 20 time points would meet the need.

Once the general timing of the peak was determined, enough animals could be studied at 3 time points to meet the second objective of finding the size (and exact timing) of the peak. With an estimated group size for this experiment of 20, 60 animals would be needed for the time point measurements, with 20 for day 0 controls and 20 for controls at the middle time point — 100 animals in all. Adding the 20 to 60 used to find the general timing of the peak still gives a substantial reduction from the original proposal.

Better-defined objectives example 3: recognising when to use a threshold approach

In this example, the objective was to screen compounds for ability to reduce skin damage from therapeutic irradiation. A standardised skin irradiation model would be used, and, based on the known variance of the model, the group size estimate was ten. Test groups for up to five compounds would be compared against a single concurrent control group. (The concurrent control was considered to be needed as differences in the absolute values of the control means could occur in different runs, though the coefficient of variation was much the same.) For 20 compounds, this would use 240 animals.

If the objective was stated slightly differently as, "to determine whether a particular compound may produce useful reduction in radiation damage", it shifts the emphasis to obtaining a qualitative result. Does, or does not, a compound have an effect above a certain threshold that would make it worthwhile to undertake further studies? In this case, a suitable threshold could be the 95% confidence level of the controls. A control group of 20 animals might be needed to get the 95% confidence level, but 20 compounds could be tested alongside, each in only two or three animals. If damage on both (or all three) of the animals tested is outside the 95% level threshold, that compound would be investigated further. If the result is equivocal, the compound could be included in the next run. Even with half the compounds tried in two runs and three rather than two replicates, only 120 animals (80 plus 40) would be needed for 20 compounds with this approach, compared to the 240 estimated above. Further reduction would be possible if it were not necessary to have a large concurrent control group for every run.

Better-defined objectives example 4: using a multifactorial approach

In the last example, the objective was to test influences of sex, age and environmental enrichment on the effect of a particular drug. Put this way it suggested a series of separate two-group comparisons — male/female, old/young, enriched/unenriched — and this was what was proposed. On the basis of variance seen in similar studies, the group size estimated was ten, giving 60 animals in all. What was really intended was to find the optimal conditions for a series of experiments with the drug and others of the same family. The objective for the initial experiment could be better put as — to determine whether sex, age and environmental enrichment are important factors in the drug effect seen.

This suggests a two by three factorial experimental design, for which eight groups of three animals should be sufficient — 24 animals in all. As well as showing whether any factor was important, the results should show any significant interactions between factors, and/or point to where further optimisation could be looked for. For a more detailed discussion of the use of a multifactorial approach see Shaw *et al*. (5).

Choice of experimental approach

As with the last two examples, rethinking the objectives may identify scope for more efficient experimental designs than the two-group comparison frequently used. These include "blocked" designs, Latin square, repeat measure, and sequential designs. A good summary of when these could be used, and of the advantages and disadvantages of each is given in Table 4.1 of the booklet, *The Design of Animal Experiments* (6). A statistician can help with avoiding problems in the more complex designs, but unless the experimenter has thought clearly about the objectives, the statistician's advice may be misplaced, as in example 2 above. As the examples illustrate, such clarity of thought cannot be assumed, and worthwhile reduction in animal use (and experimenter resources also) can be achieved by more careful specification of objectives beforehand.

Conclusion

The examples show there is scope in different types of experimental programmes for reduction, as well as refinement, both in the design of the programme (the research strategy) and in the design of the individual experiments, by taking care to define the objectives clearly.

References

1. Home Office (2000). *Guidance on the Operation of the Animals (Scientific Procedures) Act 1986*, Appendix A, p. 56. London, UK: HMSO.
2. Home Office (2000). *Guidance on the Operation of the Animals (Scientific Procedures) Act 1986*, Appendix D, p. 78. London, UK: HMSO.
3. Fry, D. & Morton, D. (1999). Non-statistical experimental design as an aid to refinement. In *Progress in the Reduction, Refinement and Replacement of Animal Experimentation* (ed. M. Balls, A-M. van Zeller & M.E. Halder), pp. 1687–1690. Amsterdam, The Netherlands: Elsevier Science B.V.
4. Fry, D.J. (1998). Relating criteria for humane endpoints to objectives. In *Humane Endpoints in Animal Experiments for Biomedical Research* (ed. C.F.M. Hendriksen & D. Morton), pp. 54–57. London, UK: Royal Society of Medicine Press Limited.
5. Shaw, R., Festing, M., Peers, I. & Furlong, L. (2002). Use of factorial designs to optimize animal experiments and reduce animal use. *ILAR Journal* **43**, 223–232.
6. Festing, M.F.W., Overend, P., Gaines Das, R., Cortina Borja, M. & Berdoy, M. (2002). *The Design of Animal Experiments: Reducing the Use of Animals in Research Through Better Experimental Design*, p. 39. London, UK: Royal Society of Medicine Press Limited.