GUIDELINES ON THE CARE OF LABORATORY ANIMALS AND THEIR USE FOR SCIENTIFIC PURPOSES

IV

PLANNING AND DESIGN OF EXPERIMENTS

LASA

LABORATORY ANIMALS SCIENCE ASSOCIATION

UFAW

UNIVERSITIES FEDERATION FOR ANIMAL WELFARE

GUIDELINES ON THE CARE OF LABORATORY ANIMALS AND THEIR USE FOR SCIENTIFIC PURPOSES

IV

PLANNING AND DESIGN OF EXPERIMENTS

LABORATORY ANIMALS SCIENCE ASSOCIATION
20 Queensberry Place, London SW7 2DZ

UNIVERSITIES FEDERATION FOR ANIMAL WELFARE
8 HAMILTON Close, South Mimms, Potters Bar, Herts EN6 3QD

ISBN 0 900767 68 5

Contents

1. PREFACE

This is the fourth section of published guidelines initiated by the Royal Society and the Universities Federation for Animal Welfare (UFAW) following the original proposal by the Laboratory Animal Science Association (LASA). The series is intended to supplement the *Guidance on the Operation of the Animals (Scientific Procedures) Act 1986* (Home Office 1990).

The first section, which dealt with the housing and care of laboratory animals, has now been superseded by the *Code of Practice for the Housing and Care of Animals used in Scientific Procedures* (Home Office 1989). The second and third sections dealt with pain, analgesia and anaesthesia (UFAW 1989) and surgical procedures (LASA/UFAW 1989). This latest section of guidelines, published jointly by LASA and UFAW, sets out some of the general principles underlying the planning and design of experimental and other scientific procedures, including approaches to statistical analysis.

The *Animals (Scientific Procedures) Act 1986*, which is in conformity with the European Council Directive (Council of European Communities 1986), provides the legal framework in the United Kingdom for the protection of vertebrate animals used for any experimental or other scientific procedure which may have the effect of causing pain, suffering, distress or lasting harm. Within the control system of project and personal licences and the designation of premises, responsibility for the proper treatment of protected animals rests primarily with the personal licensees, but also with the animal technicians and other members of staff who look after the animals before, during and after the regulated procedures. It is the prime responsibility of the project licensee to ensure that the experimental or other scientific procedure is worthwhile and that it is properly performed. Careful and thorough planning and design are a pre-requisite for success. It is essential that the project licensee has the status, qualifications, training and experience to be able to take this responsibility and to supervise and manage the project.

These guidelines on planning and design, based on a consensus of views obtained from experimenters and statisticians, should be regarded as recommendations to be used with discretion as guidance to practices and standards. All involved should conscientiously strive to achieve these standards. They are not intended as a mandatory code of practice, although where the term 'must' is used, the reader is being reminded of legal or other commonly accepted obligations. The term 'should' is generally adopted and is used to encourage attainment of desired standards.

UFAW and LASA wish to thank the principal author and those experts who helped prepare these guidelines. It is the hope and intention that the contents will assist project and personal licensees and others working under the authority of the 1986 Act in meeting their obligations and will thereby advance the welfare of the animals involved. It is particularly important in the use of animals for experimental and other scientific procedures that good science and animal welfare should continue to go hand in hand.

2. INTRODUCTION

Most of these guidelines are directed to the types of procedures in which the responses of whole animals to various treatments are compared. Such experiments are a necessary part of fundamental research and of the development and safety testing of medicinal, industrial and other products. The majority of the considerations arise from the fact that many animals may have to be used in order to obtain adequate data, to control for experimental variation, and because the range of effects of treatments and the magnitude needed to produce various levels of response are usually not known in advance.

Investigations may last for a major part of the whole of the animal's life span, which means that the effects of ageing and intercurrent disorders are likely to become important. The situation is different for many experiments performed in physiological and similar laboratories, when the effects sought are often short-lived and reversible, so that comparisons can readily be made between successive responses during simple types of experiment.

In every case the fundamental considerations are that the problem to be investigated is worth solving and that it is necessary to use the animals to do so. The experimenter must ensure that:

- The objective, purpose and likely value of the proposed study have been clearly defined.
- There has been a thorough survey of the literature, the most suitable species has been chosen, and the best experimental techniques and most appropriate dose range or treatment condition have been determined.

Consideration should be given to any need to include a pilot study to check that the proposed design will be effective, and to arranging a programme of experiments, so that less stressful investigations are done first in case they provide the answer or reduce the number of animals eventually required.

The essence of good planning is the application of the three principles of humane experimental technique classified by Russell and Burch in 1959 and known as the three Rs.

- *Replacement* of conscious living vertebrates by non-sentient material, eg tissue cultures, isolated organs, bacteria.
- *Reduction* in the number of animals required for a given degree of precision by the proper statistical design and/or the use of genetically homogeneous stock.

- *Refinement* by selection of technique so as to reduce to a minimum any pain or discomfort that may be entailed.

One of the most important features of the 1986 Act is the requirement to assess the likely adverse effects on the animals, ie the severity of the pain, suffering, distress or lasting harm, against the benefit likely to accrue as a result of the work. How this requirement is met is set out in detail in the Home Office guidance (Home Office 1990), and useful guidelines on the recognition and assessment of pain in animals have been drawn up by a Working Party of the Association of Veterinary Teachers and Research Workers (AVTRW/UFAW 1989). Evaluation and control of the severity of scientific procedures are well considered in LASA (1990). A crucial theme whenever experiments are planned is to find means to minimise the severity of the procedure by selection of the kindest techniques, by close attention to the relief of pain and distress, and by adoption of earlier and more humane end-points. Collaboration between the project and personal licensees, the Named Veterinary Surgeon, the animal technicians and others responsible for the care of the animals is very important. Prior consultation with the Home Office Inspectorate is always desirable on animal welfare grounds and may prevent wasted effort.

3. PLANNING

General

The need to use living animals as opposed to non-sentient systems must always be carefully weighed when considering experimental work, and such consideration is specifically required by the 1986 Act.

In general, experiments should not be made on living animals unless it is impracticable to achieve the intended results by alternative methods. It is recognised however that there are some types of investigation where no satisfactory alternatives are available, for example studies of:

1. Physiological processes involving interactions between different organs in the body.
2. Animal behaviour.
3. Therapeutic procedures against diseases which cannot be reproduced *in vitro*.
4. Toxicity testing when *in vitro* procedures are inadequate.
5. Aspects of disease including host-parasite relations.

Experimenters have a duty to search for methods which have the least potential for causing harm.

Selection of species and strain

If living animals have to be used, the species of choice should be the one lowest on the phylogenetic scale which is likely to satisfy the objective of the experiment. The 1986 Act refers at present to regulated procedures on vertebrates, but it allows for similar protection to be extended in future to invertebrates. Although the principles of planning and design in these guidelines are generally in relation to mammals, they may be applicable to other vertebrate classes. Moreover, it would be right to treat invertebrates in a way comparable to that which is appropriate for vertebrates, since an attitude of concern for animals of any kind should be shown at all times.

The majority of experiments are performed on mammals because of their physiological relationship to man. From this point of view it might suggest the best to use would be primates, but they should only be chosen when it is necessary to investigate features not shared by other mammals. A species with less highly developed behavioural patterns should be preferred to one in which they are more complex.

In some instances practical considerations, such as the availability, tractability and cost of a particular species or strain, may influence the choice. The availability of appropriate housing facilities, eg highly controlled environments, may also be an important factor. There are

many advantages in selecting a strain on the basis of its proven suitability, eg susceptibility to a particular disease, or response to a particular treatment. The existence of a large body of knowledge about a particular species or strain is a further valid reason in favour of choosing it. The extent of background knowledge from previous investigations, and of spontaneous disorders, is relevant, but it is important to appreciate that physiological and pathological responses may be influenced by an animal's genetic make-up and by seemingly minor factors in its environment. These may result in variation both in the results of otherwise comparable experiments, and in the findings reported by different laboratories.

In experiments where a naturally-occurring condition resembling a human disease is to be studied, the species and strain chosen will usually be restricted to those in which the condition occurs with high frequency. A valuable source for identifying such species and strains is the bibliography published by Bustad et al (1975). A more recent reference covering some of this ground is Gartner et al (1982).

More general factors which govern the choice of species in toxicity experiments are well discussed by Calabrese (1982), Tardiff and Rodricks (1987) and Roloff (1987). Much valuable information on inbred strains and the variability of their responses to treatments is given by Festing (1979). Detailed discussion of the choice of animal for testing carcinogenicity is given in a monograph produced by the International Agency for Research on Cancer (IARC 1980, Peto et al 1980) and in the discussions in Grice and Ciminera (1988); these books will also be found useful when planning prolonged experiments for other purposes.

Laboratory facilities

The number of animals required will vary in different experiments. More complex studies are usually carried out in establishments where such work is done routinely and the necessary personal skills and facilities are made available as a matter of course. However, if an experiment is contemplated in a laboratory where staff do not have the relevant experience, it is essential to make sure that the special training and facilities required at each stage of the work will be available by the time they are needed. Important aspects of the housing and care of experimental animals are dealt with in the Home Office Code of Practice (Home Office 1989). Additional points that should be specifically checked include:

1. Availability of sufficient numbers of healthy animals of the selected species, strain, sex and age, or of animals affected by the spontaneous disorder to be studied.

2. Availability of the necessary space and caging in the animal house throughout the experiment.
3. Appropriate conditions in the animal house so that the well-being of the animals will be properly maintained throughout the experiment. In some cases it will be necessary for the animals to be segregated.
4. The Animal House Manager, members of staff who will be involved, and the Named Veterinary Surgeon must be fully aware of what will be required of them.
5. A reliable and accurate system for recording, and as far as possible for checking, all the variables of interest throughout the experiment, as well as others appropriate as indicators of the health of the animals.

Although the individual scientist has ultimate responsibility, it is usually members of the animal house staff, who may not themselves hold a personal licence, who undertake animal care. As described in the 1986 Act, they will include the Named Person responsible for day-to-day care, and the Named Veterinary Surgeon to whom there must be ready access for advice and assistance. There should be close liaison between these people and the scientist during the planning stage, particularly if surgical or other more stressful procedures are envisaged.

4. DESIGN OF EXPERIMENTS

An experiment should be conducted in a manner that will maintain the highest standard of animal welfare and will provide unbiased information. Constant monitoring and self-criticism by the experimenter are necessary to ensure that this is the case. Quality control or clerical audit may be required by regulations.

For most experiments, especially the larger, more prolonged or more complex types, and those in which physiological variation or environmental factors are known to be capable of influencing the outcome, it is important to design the study in such a way as to gain the largest amount of information of the greatest reliability in an efficient manner. This will usually require early consultation with a statistician about all aspects of the design and performance of the study, in order to ensure that the experimental plan is effective and economical, and that allowance can be made for unwanted sources of variation and error. It is important to ensure that studies for regulatory purposes meet the specific requirements of as many authorities as possible.

Appropriate expert advice should be obtained as early as possible in drafting the experimental plan. It is necessary to ensure that suitable controls are employed to allow for the many and sometimes subtle ways in which unwanted variation and diverse other sources of spontaneous and acquired errors may influence the results of investigations, especially when complex or prolonged.

The conclusions of the planning stage will form the experimental protocol, which must be comprehensive and sufficiently detailed for all involved to be able to use it as a concise statement of the objectives of the experiment, as a complete guide to their responsibilities, to the husbandry of the animals used, the work to be done, observations to be recorded, how they are to be analysed and the actions to be taken in any likely eventuality. In appropriate circumstances, it will also have to contain details required to meet legal and other regulatory controls.

Biologists should consider in advance whether they are sufficiently conversant with statistical science not to need professional advice, especially for complex experiments and those involving indirect effects or interactions. If they do require help, they should ensure that they understand enough of the principles of experimental design and analysis to know what sorts of information are required by a statistician. They should be able to appreciate the arguments used by the latter and be able to weigh the value of the increased precision or sensitivity that may be offered by using more animals against the increased cost in terms of animal welfare and the human and other resources employed. The biologist and statistician together may have to weigh the advantages of designs in which fewer animals receive several treatments against the need to ensure that no animal is exposed to unacceptably repeated use.

They should also balance the attraction of a robust, simple design against the special value of more complicated types of experiment, in which there may be a risk of extensive loss of information if any part of the study fails.

Salsburg (1987) gives a down-to-earth account of one statistician's view of the design and analysis of experiments done for regulatory purposes, which affords a valuable contrast to detailed monographs on statistics. Further aspects in carcinogenicity testing are reviewed in Gart et al (1986) and Grice and Ciminera (1988).

Not all experiments involving laboratory animals require sophisticated statistical analysis, eg pilot studies, or pioneering work to elucidate basic mechanisms. If such initial work progresses to the large-scale determination of effects, or the study of toxicity over an extended period of time, it is essential to use an appropriate design to achieve the objective of the experiment efficiently and effectively. In

certain tightly controlled types of experiments, such as bioassay for pharmacopoeial purposes, the methods of design and analysis may be defined by the regulatory or scientific requirements.

Seeking a statistician's advice on methods of analysis of results after completion of an experiment is inadequate, as it is likely to be too late to compensate for faults in the experimental plan.

Type of design

When the purpose of an experiment in conceived, a statistician can give advice on the most suitable type of design and the number of animals required to achieve a given power or precision. The design will depend upon whether a single effect is being sought, a graded response, a range of effects, or changes against time, and on the need to control for environmental and other factors that influence between-animal and within-animal variation. The objective of the statistician's contribution will be to maximise the amount and reliability of the information obtained, to minimise the number of animals required to obtain an appropriate degree of precision, and to identify and control sources of variability.

Typical designs include parallel group designs, in which each animal receives only one experimental treatment; crossover designs, in which each animal receives different treatments in successive periods, and the more sophisticated factorial designs, in which more than one treatment may be examined. If, as is common, variation within-animals is less than between-animals, then for the same number of animals the treatment comparisons obtained with a crossover design are more precise than with a parallel group design. The converse applies if between-animal variation is less important, eg if strictly inbred animals are being used. In a factorial design, which may be of crossover or parallel group type, treatment comparisons can be made both between- and within-animals, but at the cost of a more complex experimental plan.

Crossover and factorial designs may make it feasible to investigate two or more treatments in the same number of animals needed to examine a single treatment in a parallel group design but with the same precision.

Crossover designs may be complete, in which each group receives each of the treatments over different, equal periods of the experiment, or incomplete, in which each group receives a different subset of the treatments being investigated. A commonly used complete crossover design is the Latin square, in which not only does each group receive each of the treatments in successive periods but also, during each of the periods, each of the treatments is being given to one, and only one group,

so that any progressive change during the experiment will affect all the treatments equally.

Problems with the crossover design are that a condition or disease may vary over a period of time and affect the results of later treatments, and that the effect of a treatment used in one period may be carried over to the next period and affect its results. Such carry-over effects may be prevented or reduced by inserting periods of no treatment or placebo treatment, or by using more elaborate crossover designs (Abeyaserera & Curnow 1984; Ebbutt 1984).

Factorial designs have the following advantages over crossover designs, as well as sharing the advantage of needing fewer animals than equivalent series of single-treatment trials:

1. The necessary duration may be less than a crossover design.
2. Each treatment is tested over a wider range of background conditions (eg presence and absence of the other treatments).
3. If the effects of the treatments are not simply additive this will become evident in the results.

Fractional factorial designs may provide as much information as crossover experiments under certain circumstances but with a smaller requirement for animals.

Factorial designs cannot be used with combinations of incompatible treatments: but in the many situations in which they are applicable it may be possible to obtain information more efficiently from a given group of animals. Their particular strength lies in the great precision that can be attained provided that the different treatments do not interact. This is especially important if it is necessary to use rare or expensive species, animals which are particularly sensitive to the experimental circumstances, or strains with inherited diseases. A useful discussion of factorial studies is given by Armitage (1971).

Sequential experimental designs have proved valuable in certain types of clinical trials (Armitage 1971), but the time required and other problems have discouraged their adoption in most other studies.

Controls

Controls may be of at least two types: negative controls representing the 'normal' or untreated state, and positive controls, treated to produce an effect, both for comparison and to confirm the adequacy of the detection system. Designating satisfactory controls is one of the most important parts of sound experimental work. An adequate negative control group of animals should be separated at random from the treatment group and then handled and housed in an identical way. In some work the controls

should receive placebo compound or the solvent vehicle. In other work the controls may have a disease condition which is not treated with the compounds under test, so that the progress of the illness can be followed to a predetermined end-point chosen to minimise distress while satisfying the experimental objectives. Positive controls treated with an active substance are sometimes necessary to provide a valid comparator.

Number of animals

The number of animals used should always be minimised, whether or not appreciable suffering is involved, and even though a considerable proportion of regulated procedures involve no suffering, eg experiments under total anaesthesia without recovery. Experimenters should ensure that the number used is necessary for the purpose of the work. For this reason there must be careful consideration of the statistical techniques needed for making this judgement.

Whatever experimental plan is adopted, one of the most important decisions to be taken before the experiment is begun is how many animals should be used. In all cases this decision requires a balance between the need for high precision or power entailing an increased number of animals, and considerations of practicability, minimisation of animal distress, and cost. It may well be that the precision or power initially considered is unacceptable, because it would require a number of animals or level of suffering that is precluded on these grounds, and in such a case a compromise has to be worked out.

When the object of the experiment is to obtain an *estimate* of a certain quantity, eg the median effective dose of the substance under investigation, it is possible to calculate the number of animals needed in order to achieve any particular degree of *precision*, usually expressed as the standard error, either in absolute terms or as a percentage of the estimate itself. For this calculation, it is necessary to know the variability of the measurement and in most cases this will have to be estimated on the basis of past experience including pilot studies. The variability concerned may be within- or between-animals, depending on the experimental design and the method of statistical analysis.

When the object is to *detect* an effect, eg a reduction in the mortality of an experimental infection, due to a particular treatment, the relevant calculation gives the number of animals needed in order to achieve any particular *power* in the experiment. Power is defined as the probability of detecting an effect at a specified level of statistical significance, if in reality an effect of given magnitude does exist. Again, a prior value for the variability must be known or assumed, and in this case it is necessary also to specify the magnitude of the effect that it is desired to detect. This

will vary widely according to the context in which the experiment is performed. Formulae and tables of power are provided, for example in Casagrande et al (1978), Fleiss et al (1980) and Lachin (1981).

Randomisation

The allocation of individual animals to particular groups should always be randomised by an objective procedure, eg use of a table of random numbers, which selects without any regard to their individual characteristics. Randomisation ensures several benefits in an experiment, most notably by preventing any bias due to conscious or unconscious selection by the investigator when deciding which animals are to receive a treatment and which are to be controls.

Randomisation is necessary for the application of meaningful statistical tests. It reduces imbalance between the treatment groups and controls as regards nuisance variables (sometimes known as confounding variables, prognostic factors or covariates) such as the age, weight and sex of the animals. It also ensures that their contribution to the uncertainty of the final result is correctly estimated and allowed for.

Randomisation alone may provide an adequate balance of nuisance factors. In big experiments the statistical law of large numbers reduces the chance of serious imbalance, but there is often particular advantage to be gained from adoption of two further steps, stratification and blocking, especially if the nature of the experiment and the interfering factors make randomisation difficult to achieve.

Stratification requires classification of the animals according to one or more interfering variables, eg body weight, with even distribution of these classes between experimental groups. Blocking involves grouping of animals according to combinations of the variables considered likely to affect the response being studied, eg body weight and sex, and random assignment of treatments within each block. Comparison of the subgroups will show the importance of the factors involved.

For a discussion of blocking and other techniques see Louis et al (1982). Professional advice is needed here in choosing the appropriate design and in analysing the results obtained.

Housing

Animals may be housed in cages, either singly or in small groups, depending on various factors such as species, age and treatment. The sex distribution, weight of the animals at the start of the experiment and the length of the experiment will govern the number of animals allocated to each cage, up to the recommended stocking density, and hence the

number of cages required. In most studies, except for certain fetal data, male and female animals are considered separately.

The physical conditions of the environment throughout experimental rooms should be uniform, as otherwise cage position may greatly influence results, particularly in prolonged or complex studies. Cage elevation, ie proximity to lights and the source of the airflow, may affect certain species of animals. It is desirable, therefore, that blocking is employed in deciding the allocation of cages in racks. The first blocking factor would be cage elevation and the second, cage position in the rack. The second factor depends upon how close to uniformity are the room conditions. If the block size were a simple multiple of the number of experimental groups, then a Latin square design could be employed. If elevation were likely to be the only important influence, the experimental design could be simplified.

Animals

The availability of suitable animals in sufficient numbers must be confirmed in the design phase. This is particularly important for inbred strains and for all larger animals. Laboratory animals can be legally obtained only from designated breeders or suppliers. They should be transported and delivered in such a way as to minimise the stress of the change in environment.

To maintain a uniform response and to minimise the risk of introducing disease, change in the source of animals during an experiment should generally be avoided. However, if health and microbiological quality can be maintained, the use of animals of different genetic backgrounds or from different sources in certain experiments may give valuable additional information. In any case, sources of animals should be recorded. For some techniques prior conditioning may be needed, or surgical preparation undertaken. Care should be taken to establish at what point the experimental treatment and observations should be commenced.

Veterinary care

There is an overriding requirement to ensure that avoidable suffering is prevented. Medication or other treatment, prophylactic or therapeutic, must be allowed, even during the experiment. Administration should be carried out under qualified and experienced supervision and details of treatment recorded. It may even be necessary to withdraw an animal from an experiment because of its need for treatment.

5. RESULTS

Analysis

How results are to be analysed should largely be decided at the stage of choosing the experimental design. A simple account of the findings may be all that is required to answer the question posed in an experiment, although the precision of every value should be considered. More refined statistical procedures are sometimes required or may be appropriate and can give much more information; guidance to a selection of reference sources is given below, but an expert opinion about their suitability should normally be obtained. Significance testing has in the past received much attention at the expense of equally simple estimation methods, such as confidence limits (Pocock 1985). The basis and limitations of any computer program should be clarified to prevent erroneous assumptions about the power and accuracy of the results. Here, too, professional advice is likely to be needed about appropriate techniques and precautions.

A number of statistical procedures of greater or lesser sophistication is mentioned as a reminder of the range of possibilities available and not to recommend one technique over another. The crucial point is that analysis of any set of data must reflect the question(s) posed, the design of the experiment and any further structure suggested by the data themselves.

Analyses can be carried out easily on a single-animal basis. Comparisons between litter mates, cage mates etc can be complicated, because of the possibility of non-independence of results (Haseman & Hogan 1975; Healey 1972; Riley & Meyer 1984). Many of these factors are best handled in the initial choice of a suitable experimental design, as when blocking or similar techniques are employed.

Simple analysis of variance and co-variance is available to test differences between groups, and the latter can be used to make some allowance for any pre-treatment group differences. More complex statistical techniques are available for continuous data, including repeated measures analysis of variance and co-variance. In these analyses, the variable of interest is the average within the response of the different groups over time and the variation of the groups over time (Winer 1971). The *treatment x time* interaction can be investigated to look for linear trends, eg average 'growth' rate, and quadratic trends, such as rate of change of growth rate. When more than two groups are involved, dose-response profiles can be tested using, for example, Williams' test (Williams 1972), or Dunnett's test (Dunnett 1964).

In analysis of variance and co-variance the underlying assumption of equality of variance should be checked, eg by Levene's test (Brown &

Forsythe 1974). If there is evidence of inequality a suitable stabilising transformation (Armitage 1971) may be applied, or a non-parametric procedure used, eg the Kruskal-Wallis one-way analysis of variance (Hollander & Wolfe 1973) with Shirley's test (Shirley 1977) to compare treatment groups with the control. When there are only two groups the t-test is often used instead of analysis of variance (to which it is then formally equivalent) and likewise the Wilcoxon rank sum non-parametric test instead of the Kruskal-Wallis test. The number and variety of analytical procedures available, the complexity of their limitations and applicability, and the potential value of subgroup or component analyses all show the need for skilled advice in the analysis of experimental results.

When the experimental unit is the litter rather than the single animal, variation in litter size can be examined by weighted analysis of variance (Armitage 1971) or, if necessary, hierarchical analysis of variance (Brownlee 1965; Healey 1972), followed by Williams' test.

When data are counts, Kruskal-Wallis one-way analysis of variance followed by Shirley's test or, for two groups, the Wilcoxon rank sum test are applicable. For proportions, the chi-square test incorporating a test for linear trend (Armitage 1971; Fleiss 1973) or Bartholomew's test (Fleiss 1973) are available. Two by two contingency tables may be analysed using the chi-square test with continuity correction or Fisher's Exact Test (Everitt 1977).

Excellent guidelines on statistical tests for all aspects of carcinogenicity and survival are clearly explained in Peto et al (1980), Gart et al (1986) and Grice and Ciminera (1988).

Reporting results

It is a common failing in written reports that insufficient information is given about the nature of the animals used, their husbandry, and details of experimental techniques. Animals should be identified by species, strain and supplier and especially noting any particular mutation of hereditary condition; sex and weight should be given, together with an indication of health status, eg barrier maintained, gnotobiotic etc, and even the microbial flora involved, if appropriate. The source and composition of the diet should be stated, together with any sterilisation procedure carried out. Details should also be provided of the feeding and watering regime and bedding composition and origin. Environmental conditions should be listed, including room temperature range, humidity range, air changes per hour, and light cycles and intensity. Details of the experimental methods employed, including anaesthesia and euthanasia, should be included. A statistical methodology section in the final report

should explain the techniques used, the assumptions upon which they were based, any computer program employed and the implication of the analysis.

6. CONCLUSION

Careful planning and design of experiments are necessary to ensure the well-being of the animals used, and are an essential part of obtaining as much value as possible from the scientific procedures.

This fourth section of guidelines covers only the general principles of the planning and design of experiments and outlines approaches to statistical evaluation and reporting results. The need for expert evaluation at an early stage is emphasised, especially for larger or more complex studies. Sources of more detailed information have been indicated where appropriate. Common design problems, such as those in long-term toxicity tests (Roe & Tucker 1973), eg inadequate randomisation, unintended variation in the position of ranks, room differences, observer differences, operator differences, poor record keeping, failure to take all data into account in the analysis, should be eliminated, so that experiments are more efficiently and effectively designed and performed.

Scientists, those responsible for the day-to-day care of animals, the Home Office Inspectorate and regulatory agencies must continue to keep themselves fully aware of developments that might improve design and performance of experiments. It is hoped that these guidelines will help to produce better studies that will entail less stress for the animals and which will require fewer of them.

REFERENCES

Abeyaserera S, Curnow R N 1984 The desirability of adjusting for residual effects in a crossover design. *Biometrics 40:* 1071-1078

Armitage P 1971 *Statistical Methods in Medical Research.* Blackwell: Oxford

AVTRW/UFAW 1989 *Guidelines for the Recognition and Assessment of Pain in Animals.* Prepared by a Working Party of the Association of Veterinary Teachers and Research Workers. Universities Federation for Animal Welfare: Potters Bar

Brown M B, Forsythe A B 1974 Robust tests for the equality of variances. *Journal American Statistical Association 69:* 364-367

Brownlee K A 1965 *Statistical Theory and Methodology in Science and Engineering.* Wiley: New York

Bustad L K, Hegreberg G A, Padgett G A 1975 *Naturally Occurring Animal Models of Human Disease: a bibliography.* Institute of Laboratory Animal Resources, National Academy of Sciences: Washington DC

Calabrese E J 1982 *Principles of Animal Extrapolation.* John Wiley and Sons Ltd: London

Casagrande J T, Pike M C, Smith P G 1978 The power function of the 'exact' test for comparing two binomial distributions. *Applied Statistics 27:* 176-180

Council of European Communities 1986 *Council Directive on the approximation of laws, regulations and administrative provisions of the Member States regarding the protection of animals used for experimental and other scientific purposes.* 86/609/EEC Council of European Communities: Brussels/Her Majesty's Stationery Office: London

Dunnett C W 1964 New tables for multiple comparisons with a control. *Biometrics 20:* 482-491

Ebbutt A F 1984 Three period crossover designs for two treatments. *Biometrics 40:* 219-224

Everitt B S 1977 *The Analysis of Contingency Tables.* Chapman and Hall: London

Festing M F W 1979 *Inbred Strains in Biomedical Research.* Macmillan: London

Fleiss J L 1973 *Statistical Methods for Rates and Proportions.* Wiley: New York

Fleiss J L, Tytun A, Ury H K 1980 A simple approximation for calculating sample sizes for comparing independent proportions. *Biometrics 36:* 343-346

Gart J J, Krewski D, Lee P N, Tarone R E, Wahrendorf J 1986 *Statistical Methods in Cancer Research, Vol III: The design and analysis of long-term animal experiments.* IARC Scientific Publications Series No 79. International Agency for Research on Cancer: Lyon

Gartner K, Hackbarth H, Stolte H (eds) 1982 *Experimental Biology and Medicine, Vol 7: Research animals and concepts of applicability to clinical medicine.* A Karger: Basel; John Wiley and Sons Ltd: Chichester

Grice H, Ciminera J 1988 *Carcinogenicity Testing. The Design, Analyses and Interpretation of Long-term Animal Studies.* Springer-Verlag: New York

Haseman J K, Hogan M D 1975 Selection of the experimental unit in teratology studies. *Teratology 12:* 165-172

Healey M J R 1972 Animal litters as experimental units. *Journal Royal Statistical Society 21:* 155-159

Hollander M, Wolfe D A 1973 *Nonparametric Statistical Methods.* Wiley: New York

Home Office 1989 *Code of Practice for the Housing and Care of Animals used in Scientific Procedures.* Her Majesty's Stationery Office: London

Home Office 1990 *Guidance on the Operation of the Animals (Scientific Procedures) Act 1986.* Her Majesty's Stationery Office: London

IARC 1980 *Long-term and Short-term Screening Assays for Carcinogens: A critical appraisal.* IARC Monographs on the evaluation of the carcinogenic risk of chemicals to humans (Supplement 2). International Agency for Research on Cancer: Lyon

Lachin J M 1981 Introduction to sample size determination and power analysis for clinical trials. *Controlled Clinical Trials 2:* 93-113

LASA 1990 The assessment and control of the severity of scientific procedures on laboratory animals. Report of the Laboratory Animal Science Association Working Party. *Laboratory Animals 24:* 97-130

LASA/UFAW 1989 *Guidelines on the Care of Laboratory Animals and their Use for Scientific Purposes: III - Surgical Procedures.* Laboratory Animal Science Association/Universities Federation for Animal Welfare: Potters Bar

Louis T A, Mosteller F, McPeek B 1982 Timely topics in statistical methods for clinical trials. *Annual Review Biophysica and Bioengineering 11:* 81-104

Peto R, Pike M C, Day N E, Gray R G, Lee P N, Parish S, Peto J, Richards S, Wahrendorf 1980 Guidelines for simple, sensitive significance tests for carcinogenic effects in long-term animal experiments. In: *Long-term and Short-term Screening Assays for Carcinogens: A critical appraisal.* Annex to IARC Monographs on the evaluation of the carcinogenic risk of chemicals to humans (Supplement 2). International Agency for Research on Cancer: Lyon

Pocock S J 1985 Current issues in the design and interpretation of clinical trials. *British Medical Journal* 290: 39-42

Riley E P, Meyer L S 1984 Considerations for the design, implementation and interpretation of animal models fetal alcohol effects. *Neurobehaviour, Toxicology and Teratology 6:* 97-101

Roe F J C, Tucker M J 1973 *Recent developments in the design of carcinogenicity tests on laboratory animals.* Excerpta Medica International Congress. Series No 311. pp 171-177

Roloff M V 1987 *Human Risk Assessment.* Taylor and Francis: Basingstoke

Russell W M S, Burch R L 1959 *The Principles of Humane Experimental Technique.* Methuen & Co Ltd: London

Salsburg D 1987 *Statistics for Toxicologists.* John Wiley and Sons Ltd: London

Shirley E 1977 A non-parametric equivalent of Williams' test for the comparison of several doses with a zero dose control. *Biometrics 33:* 386-389

Tardiff R G, Rodricks J V 1987 *Toxic Substances and Human Risk. Principles of Data* Extrapolation. Plenum Press: New York

UFAW 1989 *Guidelines on the Care of Laboratory Animals and their Use for Scientific Purposes: II - Pain, Analgesia and Anaesthesia.* Universities Federation for Animal Welfare: Potters Bar

Williams D A 1972 The comparison of several doses with a zero dose control. *Biometrics 28:* 519-532

Winer B J 1971 *Statistical Principles in Experimental Design.* McGraw-Hill: New York

Produced by Ennisfield Ltd
97 Rivington Street, London EC2A 3AY