Criteria for humane endpoints

J. Richmond

Chief Inspector, Animals (Scientific Procedures) Inspectorate, Home Office, 50 Queen Anne's Gate, London SW1 9AT, UK

Summary

This paper deals with the implementation of refinement from the viewpoint of a UK Home Office Inspector charged with ensuring that the Three Rs are implemented in research protocols which have been granted a project licence under the Animals (Scientific Procedures) Act 1986.

All of those involved in the use of animals for scientific or other experimental purposes have a moral and in many cases a legal obligation to minimize the suffering caused. The planning, implementation, continuous review and refinement of the endpoints applied are one instance of the practical application of 'refinement' as envisaged by Russell and Burch (1959). In reality the Three Rs are inseparable during the design, performance, analysis and publication of scientific studies. The use of animals for scientific purposes is an emotive and sensitive issue which can only be justified when the objectives are worthy and attainable; when non-sentient alternatives are not available; when the science cannot be improved; and the suffering is minimized. Animal experiments must be conducted with care and compassion.

The most refined research requires team work and communication. The planning, performance and analysis of animal-based experimentation are complex activities requiring that those who commission animal studies be aware of their own limitations and seek appropriate expert advice. A wide range of expert knowledge is required to plan and implement the most appropriate humane endpoints. Endpoints must be defined before work starts, be objective and reflect the true welfare cost to the experimental subjects. The specific objectives should be determined with welfare in mind. Henry Sidgwick (1901) wrote:

'Cost [is] of two kinds, either (1) the endurance of pain, discomfort or something else undesirable, or (2) the sacrifice of something desirable, either as an end or a means.'

Many regulatory systems incorporate 'utilitarian' cost/benefit analyses requiring a preliminary assessment and demonstration that the potential benefits of the work can be expected to exceed the likely animal welfare costs. The outcome of the assessment will in many cases be a matter of judgment rather than of fact.

The moral basis for cost–benefit assessment systems is undermined if animal suffering is not minimized and benefit maximized; if the framework does not require the best science and the best welfare; if the costs, benefits and endpoints are not reviewed once work is under way to determine if the initial evidence, predictions and judgments were correct; and the assessment system, and programme of work, reviewed and adjusted accordingly. Continuous improvement, the need to review, adapt and continuously improve all aspects of the Three Rs as work progresses, is inherent in the culture of care and compassion which must pervade all animal-based research activity.

There is also a need to ensure that published work properly describes how endpoints are determined and implemented, and the welfare problems encountered. This needs to be addressed not just by scientists, but by regulatory authorities, funding bodies, editorial boards and others with an interest in, or who can influence, the welfare of animals used for scientific purposes.

Refinement

Russell and Burch (1959) defined 'refinement' as '. . . to reduce to an absolute minimum the amount of distress imposed upon those animals which are still used . . . '. They acknowledged that this is a difficult area to exploit to best effect. Refinement incorporates all measures taken to avoid, minimize, recognize and alleviate pain, suffering distress or lasting harm—or to otherwise improve the welfare and well-being of the experimental subjects. Presenting refinement as 'improving welfare' is more positive than 'minimizing suffering', but there are good reasons why refinement remains a challenging area.

Minimizing suffering

The ethical imperative is to minimize suffering rather than simply reduce the numbers of animals used. Unfortunately minimizing suffering is not always synonymous with minimizing numbers or using only lower species. More aggressive protocols, the re-use of animals, and less refined endpoints may allow scientific objectives to be met with fewer animals, but with a disproportionate increase in the animal welfare cost. Similarly, asymptomatic large-animal models may on occasion be preferred to symptomatic small-animal models. Considerable experience and judgment are required to determine the defensible strategies which best minimize the overall suffering which is likely to be caused. These complex considerations are often neglected, with public and political perceptions, and agenda being derived primarily from tables of numbers and species. One danger of 'reduction targets', is that using smaller numbers of animals on more aggressive protocols with later endpoints has the potential of actually increasing the suffering caused (Hansen *et al*. 1999).

Meaningful measures

Recognition of the abnormal state depends on an awareness of, and familiarity with, normality in the species and individual under observation. Although technology is improving, measuring even basic physiological phenomena and behaviours can be difficult and may require additional interventions that add to the welfare cost or alter the parameters being measured. It has proved difficult, other than in extreme situations, to define and agree practical, objective, quantitative measures that are indicative only of pain, suffering or distress. Different stressors produce different responses and many of the events and findings used to characterize and evaluate these also occur in other contexts.

Even identifying acute pain may not always be easy, and identifying chronic pain, where the signs can be insidious, may require specific screening procedures. Some argue that we have not yet solved the real problem of finding a relatively simple, objective, practical means of accurately grading levels of animal pain, distress and suffering which can be applied to the widest range of circumstances and procedures encountered in animal-based research.

Individual response

Pain and suffering are universal phenomena with physical and emotional components unique to the individual. The behavioural and physiological responses of humans, and other animals, to painful stimuli or distressing events are not uniform. They vary between species and individuals, and in the same subject at different times. Even within matched experimental subjects responses to standard insults varies (Scharmann 1999). What is actually experienced is influenced by factors unique to the individual rather than being determined purely by the physical insult. Animals have the necessary neural architecture to perceive pain and other forms of suffering. However our understanding of

what animals actually experience and the emotions evoked is incomplete.

Another dilemma is that research to better understand pain and suffering may necessitate the production and tolerance of the very sensations and behaviours that legal and ethical guidelines for animal research require be eliminated or minimized.

Good welfare and good science are inseparable: even seemingly minor welfare problems can produce unwanted and often unrecognized physiological or behavioural effects which compromise the quality of data collected and so flawed deductions can be made when data are analysed and reported. Refinement must feature in, and arise from, all quality research programmes and all stakeholders must make better use of their influence to encourage and publicize refinement

Humane endpoints

Humane endpoints minimize the welfare costs of justifiable animal-based research. They must be described in meaningful terms, and be recognized and acted on by those entrusted with the welfare of the animals. Observation schedules should allow the prompt detection of endpoints, and there should be no temporal separation of the detection and implementation of the endpoints. Endpoints require to be adapted and contextualized to the project, experiment, and the experimental group but are best thought of as being applied to the individual animal. They may signal when an animal should cease to be an experimental subject, when specific, symptomatic or supportive treatments should be given, or when the animal should be humanely killed.

Planning

During the planning phase the likely adverse effects of the relevant acts of commission and omission (both the immediate effects, and their consequences), and how they will be recognized, must be predicted and the mildest protocol, in terms of the total suffering likely to be produced, selected. Thought must also be given to how unforeseen outcomes will be interpreted and managed. Particular care may be required to plan for potential transient pharmacological effects not indicative of true welfare problems. The planning phase should involve consultation with other experts, including those familiar with the research methods and test materials, laboratory animal scientists and veterinary surgeons, and animal care staff. In the UK, provision for this is being made at institutional level as mandatory 'local ethical review processes' are introduced.

As insights are being gained into the likely welfare costs of procedures, research workers should again ask the fundamentally important question of whether the specific experimental objective justifies the likely minimum level of suffering which is likely to be produced, or whether the objective and methods can be adjusted to provide equally useful data at a lower welfare cost.

Evidence based

It is essential to know and do 'what is right for the animals' (Russell & Burch 1959). To do this, it is necessary to know what is meaningful to the animal. The judgment of animal well-being ultimately rests with humans, and a degree of critical anthropomorphism is inevitable. 'Critical' implies empathy tempered with objective knowledge of the species (or preferably the individual animal), its needs and its behaviours, the preceding events and the significance of the signs which may be seen. Compared with the clinical situation where prior events cannot be controlled in the research setting prior events can and should be rigorously controlled.

It is important that welfare-related endpoints are evidence based in order to: prevent the needless culling of animals whose welfare may be less compromised than believed; prevent other evidence indicative of significant suffering being disregarded; make informed judgments about the severity of different procedures and models; and to evaluate claims for more refined methodologies.

A number of 'refinement' studies have been spoiled by judgments which were essentially anthropomorphic and of little significance to the experimental subject, or by failing to understand the scope and limitations of the more objective measures used. The presumption must be that in the absence of evidence to the contrary, interventions, deprivations or pathological processes that cause pain or compromise the welfare of humans, also do so in animals.

From first principles it is to be expected that the severity of a procedure is proportional to: the degree of sentience of the experimental subject; the nature, duration, strength and frequency of the challenge; the biological systems involved and mechanisms being studied; and other factors which may aggravate or ameliorate the suffering experienced by an experimental subject. Confidence is best placed in variables which occur in an appropriate context, which progress with the severity of the insult, are predictive of the ultimate welfare or pathological outcomes, and can be controlled with appropriate specific, supportive or symptomatic treatments. Signs considered to be indicative of pain should occur in contexts where there is reason to believe pain may be present, and should abate with prompt, effective analgesic administration. Signs may be evoked by adverse effects other than pain, and it should be borne in mind that analgesics can have other direct pharmacological effects which alter clinical findings. It should be a matter of routine to provide post-operative analgesia to control pain and speed the restoration of normal behaviours, such as an increase in food and water intake, so shortening the catabolic phase. Research workers should familiarize themselves with the established principle that suffering and analgesic requirements are reduced when the initial dose of analgesic precedes the painful events. It is also important to recognize that animals may be distressed, though not in pain, and therefore manifest signs which analgesics will not alleviate. This is where improving welfare by skilled staff, high standards of husbandry and care, environmental enrichment and social housing can play major roles in improving welfare.

Severity scoring systems

Numerous severity assessment systems have been published based upon discrete or continuous independent indices of impaired welfare. Continuous variables are categorized to reflect meaningful, significant differences in levels of suffering. These systems have also been successfully adapted for use in veterinary clinical management. They provide useful means of 'categorizing' the severity of specific protocols (a central feature of the UK system of controls), acknowledge that combinations of minor clinical signs may be more significant than the occurrence of any sign in isolation, and highlight that impaired welfare may be due to factors other than the local or primary effects of the interventions. They also provide a framework for determining early signs indicative of the objective having been realized, or of a poor outcome which can then be used to set a more humane endpoint. Clinical scores can be used to evaluate possible refinements to the methodology and to identify 'severe' protocols where work on replacement or refinement might most usefully be commissioned. In addition they also encourage the introduction of standard documentation, the use of plain non-technical language with a limited range of keywords to identify, describe and record clinical findings. Staff training and communication within and between research groups is simplified.

Nevertheless there are problems. They are not a substitute for proper planning. Dedicated, disciplined, and skilled staff are required for their proper implementation and predetermined endpoints must still be set. The systems must be contextualized and adapted to reflect the research objectives, models and protocols.

Situational: objective and eventuality based

Humane endpoints are determined by ethical, welfare and scientific concerns, and must cater for the following eventualities.

(1) When the experimental objectives have been realized (or when it is recognized

that they cannot be realized), even if there is no immediate welfare problem.

(2) When experimental subjects are experiencing pain, suffering, distress or lasting harm above and beyond that which can, in the context of the individual experiment, be justified or required. Such endpoints can be invoked at a low level of suffering, not simply reserved for when animals are moribund.

(3) When some intercurrent problem has compromised the data or product being collected—even when the event may not itself have seriously compromised the well-being of the animal.

Nothing must over-ride the basic principle, enshrined in European and other domestic legislation, that animals which are suffering severe pain or severe distress which cannot be alleviated must be promptly and humanely killed.

Earliest appropriate

It is important that all research is conducted to clearly defined objectives which reflect a clear understanding of the mechanisms being studied, the quality of the data which is required, and the types of deduction which are to be made when the data are analysed and published, even when working with 'standard' animal models, or well characterized test materials.

The investigation of the early events in a disease process will permit earlier endpoints than when the objective is to study later complications. Similarly, protocols to passage tumour cells or parasites will have earlier endpoints than protocols testing novel treatments. In matching objectives to endpoints, scientists must make every effort to identify early preclinical mechanistic events which provide the required new scientific insights, and to determine the earliest changes which are predictive of the subsequent outcome. Preclinical endpoints may be set requiring little or no animal suffering. As mechanisms and the predictive value of early events are better understood, opportunities for more refined endpoints result. For example, the demonstration that for some vaccines, sero-conversion reliably equates to protection has allowed this event, rather than lethal challenge tests, to establish the efficacy and potency of batches of vaccine. In some circumstances very specific mechanistic signs may be produced indicative that the scientific objective or welfare endpoint has been reached (e.g. eye irritancy testing). In other circumstances reliance may be placed in general, less specific changes, e.g. behaviour, appearance, body weight, food or water intake, or temperature. Good practical, published examples include HID50 (hypothermia-induced dose 50) as an alternative to LD50 or PD50 studies to establish bacteriological virulence (Soothill *et al.* 1992), and the observation that rodent central nervous system tumour models often show weight loss before the onset of neurological signs.

An understanding of biological mechanisms allows appropriate symptomatic or supportive therapy to be delivered which will allow the processes of interest to continue, whilst minimizing or eliminating the suffering which otherwise would have been caused. Such treatments may include analgesic therapy, specific husbandry or environmental changes, fluid replacement, or prevention of intercurrent infection. Whilst there is often concern that such treatments might compromise the science, the same is true of the untreated secondary and tertiary effects. You do not need reason to give supportive or symptomatic treatment—you need good reason not to give it.

It is important that the scientific community actively capitalizes upon its steadily increasing specialist knowledge to continually refine test protocols. In the UK it has been gratifying to see how research groups can be made to embrace the search for new refinement opportunities when they are empowered and charged with reducing the welfare costs by finding opportunities to use their latest scientific insights to refine their methodology. A good recent example has refined tumour passage work by using new information of tumour kinetics to set earlier endpoints than previously whilst increasing the yield of viable tumour cells.

Another example is the progress made to refine screening protocols for potential anticonvulsant activity where protocols at one stage requiring mortality or full-blown clinical signs gave way to reliance on earlier behavioural changes once the inevitability of these progressing was recognized. These protocols in turn are giving way to *ex vivo* screening procedures as the mechanisms of action have become better understood.

Death should seldom, if ever, be set or accepted as a required endpoint. The scientific need for such endpoints should always be questioned. As death is often the result of secondary or tertiary changes it should always be considered if lethal endpoints are consistent with good science or could be replaced by earlier euthanasia and autopsy, or be avoided by adequate observation schedules, supportive or symptomatic treatment.

Implementation

Once appropriate endpoints have been defined, it is essential that they are properly implemented. All animals should be checked by trained staff, and the inspection schedules intensified as required to facilitate the prompt recognition and alleviation of significant welfare problems. The likely resource implications should be addressed before work starts.

Staff must be properly trained and empowered to recognize and promptly implement the endpoints, with the findings and outcomes properly documented and notified to those responsible for the design and analysis of the experimental work. Welfare is not protected by systems that require that decisions on actions to be taken require lengthy internal notifications or consultations. Senior staff must ensure they are informed of problems arising during the course of studies.

In many instances recognition of the abnormal requires considerable experience of the 'normal', and regular contact with the experimental subjects. It is commendable how often experienced animal care staff are able to identify when animals are 'not right' even in the absence of recognizable clinical signs.

The endpoints and actions required must be communicated to, and understood by, the staff involved. Both the documentation and verbal descriptors used should be couched in plain language: using forms of words which will be understood by the staff checking the animals. The descriptors used should read across to other studies at the same establishment (to simplify staff training and documentation and allow refinement to be evaluated in-house), and be meaningful to those working elsewhere in the same field of research, to allow comparison with similar work performed by others, to define best practice and further raise standards.

Review

Scientific and welfare outcomes must be kept under review. Objectives should be reviewed periodically to determine if data adequate for the purpose could be obtained by modifying the original objective or adopting more refined protocols. All instances where animals are killed in extremis or found dead should be reviewed and the endpoints and observational schedules revised as necessary. In addition to the welfare considerations, such deaths may indicate that opportunities for refinement have been missed, and may require that additional animals be used. When mortalities are encountered, or animals require to be killed on welfare grounds, records should clearly distinguish between the two cases, and means of reducing the number of animals found dead determined and implemented. All such animals should be autopsied, and the information gained reviewed to consider whether they better inform the science or identify further scope for refinement.

Continuous improvement

Every opportunity must be sought to further improve on the endpoints used. It is essential that lessons learned and the more refined methodologies developed are publicized.

Scientists and editorial boards must facilitate the publication of more refined research methods as well as the new science. Much of the refinement material which is currently published appears outside the literature with which research scientists are most familiar, and may pass without notice when published, or be overlooked when new research is planned. All stakeholders should use their influence to encourage work on refinement of endpoints. Funding bodies should consider how to require that the most refined methods are used, and encourage the development and reporting of refinements developed during the course of funded work. Regulatory authorities should strive to produce and publicize minimum severity protocols.

This meeting provides a showcase for best contemporary practice. The practical insights provided in the original work presented will produce immediate benefits through those attending the meeting or who read the proceedings.

References

Hansen AK, Sandøe P. Svendsen O, Forsman B, Thomsen P (1999) The need to refine the notion of reduction. In: *Humane Endpoints in Animal Experiments for Biomedical Research* (Hendriksen CFM, Morton DB, eds). London: Royal Society of Medicine Press, pp 139–44

Russell WMS, Burch RL (1959) *The Principles of Humane Experimental Technique*. London: Methuen & Co [Special edition, 1992, Potters Bar: UFAW]

Scharmann W (1999) Physiological and ethological aspects of the assessment of pain, distress and suffering. In: *Humane Endpoints in Animal Experiments for Biomedical Research* (Hendriksen CFM, Morton DB, eds). London: Royal Society of Medicine Press, pp 33–9

Sidgwick H (1838–1890) Quoted from *Principles of Political Economy*, 3rd edn (1901) Volume 3, Chapter 1

Soothill JS, Morton DB, Ahmad A (1992) The HID50 (hypothermia-inducing dose 50): an alternative to the LD50 for measurement of bacterial virulence. *International Journal of Experimental Pathology* **73**, 95–8