Possibilities for Refinement and Reduction: Future Improvements within Regulatory Testing

Martin L. Stephens  
*The Humane Society of the United States*

Kathleen Conlee  
*The Humane Society of the United States*

Gina Alvino  
*The Humane Society of the United States*

Andrew N. Rowan  
*The Humane Society of the United States*

Follow this and additional works at: [https://www.wellbeingintlstudiesrepository.org/acwp_arte](https://www.wellbeingintlstudiesrepository.org/acwp_arte)

Part of the Bioethics and Medical Ethics Commons, Laboratory and Basic Science Research Commons, and the Research Methods in Life Sciences Commons

**Recommended Citation**

Possibilities for Refinement and Reduction: Future Improvements Within Regulatory Testing

Martin L. Stephens, Kathleen Conlee, Gina Alvino, and Andrew N. Rowan

The Humane Society of the United States

CITATION


KEYWORDS

adverse effects, analgesia, anesthe sia, animal use alternatives, carbon dioxide, euthanasia, methods, research design, statistics, numeral data

ABSTRACT

Approaches and challenges to refining and reducing animal use in regulatory testing are reviewed. Regulatory testing accounts for the majority of animals reported in the most painful and/or distressful categories in the United States and Canada. Refinements in testing, including the use of humane endpoints, are of increasing concern. Traditional approaches to reduction (e.g., improving experimental design) are being supplemented with complementary approaches, such as the use of tier testing to eliminate some chemicals prior to in vivo testing. Technological advances in telemetry and noninvasive techniques will help decrease either the demand for animals in testing or animal suffering. Further decreases in animal use will stem from international harmonization and coordination of testing programs. Progress in refinement and reduction faces a variety of broad challenges, including limited funding for research. In the specific area of refinement, a key challenge is the issue of distress (as distinct from pain). In the area of reduction, the practice of using unjustifiably high numbers of animals from small species (e.g., rodents) should be challenged. One case study of the use of carbon dioxide as a euthanasia agent illustrates the need for further analysis and research. Notwithstanding the complexities and challenges, the potential for refinement and reduction in regulatory testing is encouraging.

Introduction

Russell and Burch (1959) proposed the framework of the 3Rs of refinement, reduction, and replacement more than 40 yr ago. Since that landmark publication, significant progress has been made, especially in the arena of regulatory testing (Stephens et al. 2001). Several reviews of refinement and reduction alternatives have been written in recent years (e.g., Festing 1999; Morton 1995, 1998; Rowan 1995). The present review concentrates on regulatory testing and focuses primarily on refinement.

We first assess animal use patterns in regulatory testing and the associated levels of animal pain and distress experienced in this testing to provide a better sense of use patterns, historical trends, and how best to deploy our resources. Perhaps the most comprehensive national statistics come from Great Britain. Of the 2.7 million animals used there in 1998, 21% were used in toxicity testing (Home Office 1999). The types of toxicity studies that account for the bulk of this testing include acute/subacute lethal
studies (124,000 animals); acute/subacute range-finding, limit testing, and related studies (109,000); and reproductive/teratogen/mutagen studies (71,400).

Data from the United States (Stephens et al. 1998) and Canada (CCAC 1998a) indicate that testing procedures account for the vast majority of the animals reported in the highest categories of pain and distress. These findings underscore the importance of refining and reducing animal use in regulatory toxicity testing.

**Approaches to Refinement**

Refinements are modifications in animal-based procedures that either decrease pain, distress, and discomfort or increase animal well-being (Morton 1995). They can be applied to all aspects of animal care and use in the laboratory and can improve scientific outcomes as well as animal welfare (Smaje et al. 1998). Some key approaches to refinement, summarized below, include humane endpoints, pain and distress relief, dosing limits, mechanism-based methods, and other approaches.

**Humane Endpoints**

Humane endpoints are refinements that allow a study to be ended early, without compromising the experiment’s objective or results, thereby precluding any further animal pain and distress. For example, Cussler and colleagues (1999) demonstrated that circling behavior in mice reliably predicts death in the rabies vaccine challenge test. Consequently, the onset of this sign can be used as a humane endpoint in lieu of death. Both the Canadian Council on Animal Care (CCAC 1998b) and the Organization for Economic Cooperation and Development (OECD 2000) have recently issued guidelines on humane endpoints. ILAR Journal recently devoted an entire issue to the subject (ILAR 2000; also available online at <www.nationalacademies.org/ilarjournal>), and the proceedings of a 1998 conference on the subject have been published (Hendriksen and Morton 1999).

**Pain and Distress Relief**

An obvious approach to refining a protocol that would otherwise cause pain or distress is to provide pain- or distress-relieving drugs such as anesthetics, analgesics, tranquilizers, and sedatives to preclude or alleviate these adverse effects. Although this approach is routine in procedures involving surgery, it may also be indicated in other protocols. Administration of test substances through gavage is a common procedure in regulatory testing. Murphy and colleagues (2001) found that brief halothane treatment before daily gavage administration of vehicle reduced stress-induced weight loss in rats. However, anesthetization was associated with incomplete vehicle retention. More troubling, the authors found a significant level of gavagerelated deaths, which anesthesia seemed to dramatically reduce (5 of 9 nonanesthetized animals died or required euthanasia vs. 1 of 37 anesthetized animals). Subacute or chronic toxicity studies using gavage could not be carried out if gavagerelated mortality were this high. It is probable that there were problems with the technical skills of those performing the gavage (see Straughan 2001).

**Dosing Limits**

Mechanism-based Methods

A better understanding of the mechanism of a disease process can sometimes enable scientists to modify or completely redesign an animal-based test procedure. For example, the traditional test for identifying substances that cause allergic contact dermatitis (ACD)\footnote{ACD} is the guinea pig maximization test, in which subjects are exposed to a substance and then monitored for adverse reactions—in this case, the clinical signs of ACD (e.g., redness, swelling, blistering). By contrast, the newer, mechanism-based test, the murine local lymph node assay, capitalizes on the finding that sensitizers induce proliferation of lymphocytes in the lymph node draining the site of chemical application (Kimber and Weisenberger 1989). Under appropriate test conditions, this lymphocyte proliferation is proportional to the applied dose and provides a means of obtaining an objective, quantitative measurement of sensitization. The mice in the murine local lymph node assay are euthanized before they show any clinical signs of ACD.

Other Refinement Approaches

There are many other approaches to refinement in regulatory testing and other animal-based procedures. These approaches include, but are not limited to, training personnel in best practices; ensuring adequate levels of staffing and monitoring; housing animals in nonbarren, even enriched, enclosures; using indwelling catheters and tethers or osmotic pumps in lieu of repeated injections; acclimatizing animals to testing apparatus or training them (particularly primates) to cooperate in chemical administration; and housing animals in groups (e.g., Healing and Smith 2000).

Approaches to Reduction

Reduction alternatives are methods that use fewer animals than the conventional procedures but still yield comparable levels of information, or methods that use the same number of animals but yield more information, so that ultimately, fewer animals are needed to complete a given project or test (Balls et al. 1995). Reduction alternatives are typically viewed as modifications of existing procedures. For purposes of this review, we adopt a broader view to encompass any changes in practice that result in fewer animals being used, without compromising test results. Some of the main approaches to reducing animal use in regulatory testing are summarized below.

Improving Experimental Design

Festing (1994) has discussed numerous ways to design experiments that use fewer animals while yielding equivalent or greater statistical power. For example, sequential designs often use substantially fewer animals than dosing all animals at the start of a procedure. This approach has been successfully applied to acute toxicity testing, with the fixed dose procedure, acute toxic class method, and the up-and-down method, each of which uses fewer animals than the conventional LD50 test (Festing et al. 1998). Using the limit test is another way to reduce animal use in acute toxicity studies.

Using Genetically Defined Animals

Russell and Burch (1959) and Festing (1999) have championed the use of animals with genetically defined back grounds as a way of limiting statistical variance and thereby reducing the number of animals needed to achieve a given level of discriminatory power. This approach includes use of in-bred strains as well as crosses of two in-bred strains.
Using Screens in Tiered Testing

Using nonanimal screening procedures in a tiered testing scheme can eliminate some chemicals from further testing in higher tiers, thereby reducing animal use. This approach has been applied to several areas, including eye irritancy (with pH determination precluding some Draize testing) and carcinogenicity (with a positive Ames test substituting for 2-yr rodent bioassays). Prescreens offer promise in the field of acute toxicity testing, and cytotoxicity tests are being explored as a means of screening out toxic chemicals, in particular (NIEHS 2001; Spielmann et al. 1999). Prescreens also offer promise in the field of endocrine disruptor testing, where quantitative structure-activity relationships and high throughput prescreens are being explored.

Technological Advances Facilitate Refinement and Reduction

A number of technological advances appear promising in the context of reducing or refining animal use in regulatory testing. Telemetry is already being applied to the field of drug discovery. However, it does entail surgical implantation of a transmitter, and it could be argued that the pain and distress associated with this surgery offsets the distress subsequently avoided through remote monitoring. Nevertheless, telemetry has the potential to reduce animal numbers dramatically (Kramer et al. 2001). The reduction stems from the ability to monitor the same animal over time repeatedly, rather than euthanizing animals at various times to assess treatment effects. To date, the application of telemetry to regulatory testing has been limited to experimental assessment of specific chemicals, primarily pharmaceutical candidates. However, the increasing miniaturization, sophistication, and affordability of telemetry will allow this technology to be used more widely in testing.

Similarly, advances in imaging techniques such as magnetic resonance imaging and positron emission tomography, like telemetry, allow the same animal to be monitored repeatedly over time, thereby reducing animal numbers (Balaban and Hampshire 2001; Cherry and Gambhir 2001; Paulus et al. 2001).

In the future, microarrays of genetic material (e.g., genomics), their products (proteomics), and metabolites (metabonomics) may serve as prescreens and, eventually, definitive tests, thereby reducing or replacing animal procedures for some testing applications. Similarly, integrated computer modeling may one day produce a "virtual human" (ORNL 2001), which could decrease animal use in regulatory testing.

Policy Issues

Regulatory testing is the outcome of policy decisions at national and international levels. These decisions have ramifications for the scale of animal use and any resulting pain and distress. In pharmaceutical testing, representatives of government and industry in Europe, the United States, and Japan saw the need for international harmonization of testing requirements and protocols and established the International Conference on Harmonization to meet this challenge. Although the mission of the conference was not animal welfare per se, its harmonization efforts have significantly reduced animal use in pharmaceutical testing (Osterberg 2001).

Similarly, there is a critical need to coordinate large-scale testing programs being developed at both national and international levels, such as efforts to assess endocrine disrupters. The targets for coordination include the chemicals and endpoints to be assessed, the protocols to be used, and validation of chosen protocols. The Organization for Economic Cooperation and Development has begun playing a role in coordinating these programs.
Animal protection organizations are beginning to be recognized as stakeholders in the development of testing programs that may involve animal use. Consequently, their voices are now beginning to be heard when government agencies formulate these programs, which should help ensure proper consideration and integration of the 3Rs.

The evolution of the US high production volume (HPV1) chemical testing program illustrates how concerns about the 3Rs and animal welfare, raised by animal protection organizations and the alternatives community, can reduce the scale of animal use in a developing program and promote the development of alternative methods, all without compromising the goals of the program. The US Environmental Protection Agency (EPA1) initially developed the HPV program without seeking any input from animal protection stakeholders. After the program was publicly announced in 1998, animal protectionists voiced their concerns to EPA, Congress, and other decision makers. The Center for Alternatives to Animal Testing (Johns Hopkins University, Baltimore, Maryland) held a series of workshops on how the 3Rs could be applied. This attention culminated in an agreement in which the EPA pledged to make several changes in the HPV program that would lead to a substantial reduction in animal use. The government also agreed to spend $4.5 million on the development of relevant nonanimal methods.

Challenges for the Future

Broad Challenges

There are several broad challenges to the development and implementation of refinement and reduction alternatives that should be addressed to expedite progress in advancing humane experimental technique. These challenges include the following:

- Encouraging greater funding of alternatives research

  We see funding as the major driver of innovation in the 3Rs. Consequently, the current, limited level of such funding is a major impediment to progress (see <http://altweb.jhsph.edu/databases/funding/funding.htm> for a country-by-country list of available sources of alternatives funding). Modest levels of funding can go a long way in refinement and reduction research, in that expensive, large-scale validation exercises are usually not needed. Those seeking to persuade funding institutions, including governments, to support alternatives research should emphasize the resulting benefits to science as well as animal welfare.

- Better ways to translate best practices into standard practices

  It is not enough simply to develop new refinement and reduction alternatives; these techniques must be incorporated into practice. Regulatory requirements to consider alternative methods can help foster the implementation of new techniques. However, an important prerequisite of progress is to publish or otherwise disseminate innovations. Professional societies should encourage the publication of new alternative techniques relevant to their disciplines, and, when appropriate, regulatory agencies should bring significant developments in alternative methods to the attention of regulated parties, as was the case with monoclonal antibody production in the United States (see <http://altweb.jhsph.edu/publications/misc/oprr.htm>).
• Developing new testing programs that incorporate refinement, reduction, and replacement alternatives into regulators’ programs to the fullest extent possible, and to avoid any bias in assessing the applicability of in vitro versus in vivo methods.

  Technical experts should work with more politically oriented proalternative organizations to ensure that regulatory agencies give proper consideration to existing alternative methods. Moreover, agencies attempting to launch large-scale animal testing programs should devote substantial funding to research and development of alternative methods, so that such methods can be incorporated into the programs.

• Harmonizing national statistics on animal use patterns so that rational priorities for reduction and refinement research can be identified internationally.

Given the limited funding currently available for alternatives research, priorities for this research should be set based on pertinent considerations such as the number of animals used in a given procedure annually and the level of pain and distress entailed in the procedure. This information can sometimes be gleaned from national statistics on animal use, but it would be even more helpful if harmonized international figures were available. The European Union has begun issuing periodic reports on the number of research animals used by member states (the latest report is available at <http://europa.en.int/comm/environment/document/99191_en.htm>). This first step toward a more global effort is promising.

Specific Challenges: Refinement and Reduction

In addition to the more general challenges identified above, several more specific impediments to progress in refinement and reduction can be identified. In these areas, we believe the following challenges are among the most important:

• Better understanding of distress, especially non-pain-induced distress and discomfort, and how to identify, quantify, and prevent or alleviate it (Rowan et al. 1998)

  Considerably more attention has been paid to pain than to distress, and we believe this imbalance has led to an underappreciation of the incidence and severity of distress, as well as to the adverse impact of any distress on the scientific results and conclusions.

• More reliable indicators of adverse effects to determine the relative roles of indices (e.g., weight loss, body condition, behavior, and hormone levels) and how multiple indicators can be combined into an integrated approach, as in the use of score sheets (e.g., Morton 2000)

  Good assessment measures are critical to evaluating potential refinements (Flecknell 1994).

• Development of endpoints well before the moribund state

  As a result of the call for more humane endpoints, death as an endpoint is increasingly questioned. The most common endpoint substituting for death is the moribund condition. Although the motives are good, an animal that reaches the moribund state probably experiences as much distress as one that dies.
• Discrimination on the basis of body size without adequate justification

When a choice between species is possible, the practice of using higher numbers of animals simply because they are smaller (or less expensive) should be challenged (Balls et al. 1995).

• Animal waste from single-sex testing requirements

Mismatches between supply and demand create surpluses of one sex or the other that result in euthanizing unwanted animals (Festing et al. 1998).

• An array of often-overlooked variables

Variables such as circadian rhythms and repeated handling not only can compromise experimental results in pharmacology and neuroscience research but also can lead to unnecessarily high levels of animal use or pain and distress (Claassen 1994). The proper control of these variables will not only improve the quality of experimental results but can also refine or reduce animal use.

Case Study: Carbon Dioxide as a Euthanasia Agent

One of the key approaches to challenging the status quo in animal use is to conduct detailed assessments of common animal-based techniques to determine whether these techniques are candidates for refinement, reduction, or replacement. The Humane Society of the United States has conducted such a review of the use of carbon dioxide as a euthanasia agent (see <http://www.hsus.org/ace/11427>). Our analysis of published data raises several concerns about the routine use of carbon dioxide, which include the following:

1. The evidence on whether C02 causes pain and distress in animals is mixed. Some C02 studies report no pain or distress (e.g., Hewett et al. 1993); others report the opposite (e.g., Coenen et al. 1995).

2. C02 has been used in human and rodent studies as a pain- and stress-inducing stimulus (e.g., Anton et al. 1992), which suggests that C02 is indeed painful or stressful to animals, at least under certain conditions.

3. Adverse reactions, including seizure, nose hemorrhage, rearing, defecation, and excessive salivation, have been noted in rodents and other species at C02 concentrations of >50% (e.g., Ambrose et al. 2000).

4. Histological analyses of animals reveal lung edema and hemorrhage at all concentrations of C02 use (e.g., Fawell et al. 1972).

5. The literature is mixed as to which of two common methods of C02 induction (prefilling the chamber vs. gradual induction) is preferred. Distress has been reported with each method.

6. The shortest time to collapse reported at the 70% C02 concentration recommended by the American Veterinary Medical Association is approximately 10 sec (Mischler et al. 1994) although one study reported that anesthesia did not occur until 4.01 min (Danneman et al. 1997). This
duration allows for a significant amount of time for the animal to suffer before becoming unconscious.

7. Humans experience pain when the CO2 concentration is 50% and higher (e.g., Danneman et al. 1997). US Government Principle #4 and US Department of Agriculture Policy #11 both state that procedures causing pain and distress in humans should be assumed to cause pain and distress in animals, absent “evidence to the contrary” (US Government Principle #4; PHS 1993). In the case of CO2, the animal data are equivocal but the human data are clear. Therefore, The Humane Society of the United States argues that the continued use of CO2 as a euthanasia agent is inconsistent with government principle #4 and US Department of Agriculture policy #11.

A number of organizations, including the American Veterinary Medical Association (AVMA 2001), the Canadian Council on Animal Care (1993), the European Commission (1996), and the Australian and New Zealand Council for the Care of Animals in Research and Teaching (Reilly 1993), currently support the use of CO2 for euthanasia of rodents and provide guidelines on its use. These guidelines differ in significant ways. In summary, the evidence of potential pain and distress associated with the use of carbon dioxide as a sole agent for euthanasia indicates that its routine use for this purpose should, at the very least, be questioned. Consequently, refinements to the use of CO2 as a sole agent for euthanasia, such as the use of an inhalation anesthetic before exposing animals to CO2, should be considered. Finally, the use of CO2 at lower concentrations as an anesthetic agent must also be questioned.

Concluding Remarks

Reduction and refinement alternatives have significant potential to decrease the use and suffering of animals used in regulatory testing further. The pace of future progress in these areas will depend on how well several challenges are met, including expanding funding for alternatives research.

Footnotes

1 Abbreviations used in this presentation: ACD, allergic contact dermatitis; EPA, US Environmental Protection Agency; HPV, high production volume.

References
