Increasing the Implementation of Alternatives to Laboratory Animal Use

Jasmijn de Boo and Andrew Knight

Animal Consultants International, London, UK

Abstract
The scientific and logistical limitations incurred by the use of animal models of humans within biomedical research and toxicity testing are substantial, and increasingly recognized; as is social concern about, and consequent regulatory restriction of, laboratory animal use. In defiance of these factors, such use remains enormous. Based on best estimates, 11,154,961 living non-human vertebrates were subjected to fundamental or medically-applied biomedical research, toxicity testing, or educational use, within Japan, in 2004; which was second only to the US. Additionally, the use of genetically-modified animals, and the implementation of large-scale chemical testing programs, are increasing laboratory animal use internationally. These trends demonstrate the need for considerably greater awareness of, and compliance with, the principles of the 3Rs—namely, the replacement, reduction and refinement of laboratory animal use—within governmental, academic and commercial sectors. These principles are widely recognized as essential to good laboratory animal practice. They may increase research quality and the robustness of outcomes, result in reduced timeframes and resource consumption, and jointly benefit consumers, industry and laboratory animals. An overview of 3Rs principles, and of strategies likely to increase their implementation, is therefore provided. Combinations of such strategies may have synergistic effects, improving both scientific outcomes and animal welfare.

Key words: 3Rs, alternative, animal experiment, animal model, laboratory animal.

Trends in laboratory animal use
Worldwide, millions of animals have been used in toxicity testing and biomedical research aimed at developing cures for human diseases. Based on publication rates, Taylor and colleagues (2008) estimated that 58.3 million living non-human vertebrates were subjected to fundamental or medically-applied biomedical research, toxicity testing, or educational use, within 179 countries, in 2005. When animals killed for the provision of experimental tissues, used to maintain genetically-modified (GM) strains, or bred for laboratory use but killed as surplus to requirements, were included, the estimate increased to a total of 115.3 million vertebrates used worldwide in 2005. However, this calculation relied on simple, arithmetic means. Use of more accurate weighted means raised the total to 126.9 million (Knight, 2008a).
Due to several factors, however, these estimates were considered to be highly conservative. As identified by Taylor and colleagues, for example, their estimate of 17.3 million living vertebrates used within the US is very significantly less than a 2000 US Animal and Plant Health Inspection Service estimate of 31–156 million, based on extrapolation from the results of a survey of only 50 of 2,000 research institutions (USDA, 2000).

Two important factors are increasing laboratory animal use internationally. Dramatic increases in the use of GM animals have recently reversed previous steady decreases in some countries. The production and maintenance of GM strains requires substantial breeding, further increasing numbers. Within the UK, for example, a steady and significant overall reduction since 1976 stabilized during the early 1990s, and then reversed. 3.1 million living vertebrates and one species of octopus were used in 2007—the highest number for around 15 years (Home Office, 2008). In 1995, GM animals were used in 8% of regulated procedures. By 2007 it was 36% (Home Office, 2008). Increased GM animal use has also been recorded in Germany (Sauer et al., 2005) and Switzerland (Anon., 2007), where total animal use is also increasing (Anon., 2007; Rusche, 2003).

Additionally, unprecedented increases in laboratory animal use have recently been proposed within several chemical testing programs, intended to rectify knowledge gaps regarding the toxicity of chemicals produced or imported into Europe or the US in particularly high quantities, or that otherwise raise special concerns (Combes et al., 2004; Green and Goldberg, 2004). Such programs follow increasing public, political and regulatory concern about the potential toxicity of a wide range of environmental, occupational and consumer chemicals.

**Scientific limitations of animal models**

Despite widespread reliance on animal models of humans within fundamental or clinically-applied research and toxicity testing, the scientific limitations incurred by such models are considerable, wide-ranging, and increasingly recognized. These may include differences between species and genders—with subsequent effects on toxicokinetics (the study of bodily distribution, particularly absorption, distribution, metabolism and excretion), or pharmacodynamics (the study of mechanisms of action, and drug effects). Others may include: the use of unrealistic doses and exposure durations; the loss of biological variability or predictivity, resulting from the use of inbred strains, young animals, restriction to single genders, and inadequate group sizes; the lack of co-morbidities or other human risk factors; and stress-related physiological or immunological distortions; among others (Hartung, 2008a-b; Matthews, 2008).

Interspecies variations in P450-dependent monoxygenases, for example, are well established (Guengerich, 2006). These constitute the major family of xenobiotic metabolising enzymes—that is, enzymes catalysing the oxidation (i.e., the metabolism) of foreign compounds, such as drugs or toxins. Their major purpose is the generation of non-toxic blood-soluble metabolites (products of metabolism, usually by the liver) suitable for renal (kidney) or other elimination. Interspecies differences in metabolic pathways, rates and products may decrease efficacy or produce toxicity, and are a key cause of high clinical trial failure rates during pharmaceutical development (DiMasi et al., 2003). In fact, only eight percent of candidate drugs which have passed animal studies are successfully licensed for human use by the US Food and Drug Administration (Crawford, 2004).

The utility of animal models has been repeatedly questioned. Of 20 published systematic reviews examining human clinical utility located during a recent, comprehensive search, animal models demonstrated significant potential to contribute toward the development of clinical interventions in only two cases, one of which was contentious. Seven additional reviews failed to demonstrate utility in reliably predicting human toxicological outcomes, such as carcinogenicity and teratogenicity. Results in animal models were frequently equivocal, or inconsistent with human outcomes (Knight, 2007).

Because such reviews examine large numbers of animal experiments or other procedures selected systematically, via randomization or similarly impartial and methodical means, they provide a high level of reliability. They have assessed utility via citation analyses (primarily), as well as detailed examination of appropriate studies individually. They demonstrate that many animal experiments are never cited by future publications, and that the

utility of such experiments within biomedical research and toxicity testing is often less than expected.

Similarly, a European working group comprised of 18 pharmaceutical companies recently conducted an evidence-based review of the value of acute toxicity studies conducted on 74 compounds. Such animal-based studies—which are usually conducted in rodents—have been required prior to the licensing of pharmaceuticals intended for human use. This is the only study type during the process in which lethality is mentioned as an endpoint (Robinson et al., 2008). However, the acute toxicity data for these 74 compounds was never used to: (i) terminate drugs from development, (ii) support dose selection for subsequent repeat-dose animal studies, or (iii) set doses, identify target organs or indicate biomarkers for monitoring in human clinical trials—a minority of which also occurred prior to animal studies. The conclusions—which were discussed and agreed with representatives of regulatory bodies from the US, Japan and Europe, in 2006—were that such studies are not necessary prior to first clinical trials in humans, and that sufficient information can be obtained from other studies conducted at more relevant doses for humans, which already comprise an integral part of drug development (Robinson et al., 2008).

**Regulatory limitations of animal models**

Increasing social concern about invasive animal use within biomedical research and toxicity testing has driven the evolution of related legislation and regulatory policies within many countries and regions (see http://www.vetmed.ucdavis.edu/Animal Alternatives/policies&regs.html for summaries). Such policies seek to restrict laboratory animal use to only those instances where non-animal alternatives are considered scientifically inadequate, for investigating questions of importance sufficient to justify animal use. Within the US, for example, the regulations associated with the Animal Welfare Act (AWA, 1966, most recently amended in 1990) require documented consideration of alternatives, “to any procedure that would reasonably be expected to cause more than slight or momentary pain or distress in a human being.” Some warm-blooded vertebrates are protected, with important exceptions—primarily birds, mice and rats (despite the fact that the latter rodent species comprise the overwhelming majority of laboratory animals). Such consideration is required whether or not analgesics or anesthetics are used, and research must not unnecessarily duplicate previous experiments (US Government Printing Office, 1985; Kulpa-Eddy, 2006; Larson, 2006).

The Japanese animal protection law was similarly amended in 2005, stressing the importance of alternatives. Associated guidelines were announced in 2006 by the Ministries of Education, Culture, Sports, Science and Technology; of Health, Labor and Welfare; and of Agriculture, Forestry and Fisheries. These guidelines confirm the responsibility of research institutes to conduct animal experiments only under ethics committee approval. Education about the 3Rs (see following) must also be provided to researchers (Ohno, 2008).

**Deficiencies in consideration of alternatives**

Resistance to the use of alternatives remains considerable, however, within some governmental, academic and commercial sectors. A 1999 UK survey found that although most animal researchers reported positive attitudes toward alternatives, nearly 20% did not support their use on moral or ethical grounds, and most did not believe that replacement alternatives could provide scientific information of equivalent value to that obtained from animal experiments (Purchase and Nedeva, 2001).

In 2000, a US survey identified a range of common areas of non-compliance of researchers with the regulations associated with the AWA. The most common was inadequate consideration of alternatives (at 600-800 research facilities), and the fourth most common was unnecessary experimental duplication (at approximately 250 facilities). Other deficiencies identified included inadequate justification for animal numbers, and alleged uncertainty of research personnel about signs indicative of pain and/or distress (USDA-APHIS-AC, 2000).

Similarly, a 2007 survey of 24 Japanese pharmaceutical institutes and universities identified deficiencies such as inadequate consideration of alternative methods, of animal pain and distress, and inadequate education of researchers about alternatives principles, methods to evaluate and decrease pain and distress, and euthanasia techniques (Ohno, 2008).

Such attitudes about laboratory animal use reveal a
marked lack of consideration of scientific constraints on the human utility of animal models; of social concerns about their use; and about the unprecedented logistical challenges that reliance on animal models will inevitably incur, during high-throughput chemical testing programs. They also demonstrate marked and widespread deficiencies in awareness of the potential and availability of non-animal methods. Accordingly, reviews of such methods, and of suitable strategies for facilitating their implementation, are urgently warranted.

3Rs principles
It is now considered fundamental to good laboratory animal practice in scientific research, product testing and other technical procedures to conduct such procedures in accordance with the ‘3Rs,’ which were first proposed by Russell and Burch in 1959. These are the:

1. Replacement of animal use with non-animal alternatives, wherever possible;
2. Reduction of animal numbers to the minimum possible;
3. Refinement of animal use, in order to avoid or minimize animal pain, distress or other adverse effects suffered at any time during the animals’ lives; and to enhance well-being (Zucco et al., 2005).

Additional Rs have occasionally been proposed by others. Re-use or Recycling of animals aims to reduce total numbers. Rehabilitation aims to ensure protection and care of animals after procedures have been terminated (Anon., 1986; Pereira and Tettamanti, 2005).

Replacement alternatives
A considerable variety of strategies may facilitate the replacement of animal use within biomedical research and toxicity testing, which we recently reviewed in detail (Knight, 2008b). These include: mechanisms to enhance the sharing and assessment of existing data prior to conducting further studies; physicochemical evaluation and computerized modeling, including the use of structure-activity relationships (SARs) and expert systems; and the use of minimally-sentient animals from lower phylogenetic orders, or early developmental vertebral stages, as well as microorganisms and higher plants.

SARs predict biological activity such as toxicity on the basis of molecular structure. Computerized expert systems seek to mimic the judgment of expert toxicologists, by using known rules about factors affecting toxicity, in combination with physicochemical or other information about a specific compound. They make predictions about toxicity and related biological outcomes, such as metabolic fate.

A variety of tissue cultures, including immortalized cell lines, embryonic and adult stem cells, and organotypic cultures, are also available. In vitro assays utilizing bacterial, yeast, protozoal, mammalian or human cell cultures exist for a wide range of toxic and other endpoints. These may be static, or perfused, and used individually, or combined within test batteries. Human hepatocyte (liver cell) cultures and metabolic activation systems offer potential assessment of metabolite activity—a very important consideration when assessing toxicity.

Microarray technology, allowing assessment of large numbers of genes simultaneously, may allow genetic expression profiling (detection of up- or down-regulation of gene activity, caused by exposure to test compounds), increasing the speed of toxin detection, well prior to more invasive endpoints.

The safety profile and predictivity for diverse patient populations of human clinical trials should be improved using microdosing, staggered dosing, more representative test populations, longer exposure periods, and other strategies. Surrogate human tissues, advanced imaging modalities, and human epidemiological, psychological and sociological studies, may all increase understanding of disease etiology (causation) and pathogenesis (development), and facilitate the development of safe and effective pharmacologic or other interventions.

Particularly when human tissues are used, non-animal models may generate considerably faster, cheaper results, more reliably predictive for humans, and may yield greater insights into human biochemical processes. In recognition of these factors, a key partnership was announced in 2008 between the US Environmental Protection Agency (EPA) and two other federal agencies. This partnership aims to meet the future toxicity testing needs of the EPA and the National Toxicology Program (NTP), by implementing recommenda-
tions proposed in the NTP’s Roadmap for the Future (NTP, 2004), and the report of the National Research Council’s Committee on Toxicity Testing and Assessment of Environmental Agents, Toxicity Testing in the 21st Century: A Vision and a Strategy (NRC, 2007) (Collins et al., 2008; Hartung and Leist, 2008; Leist et al., 2008). These reports propose significantly increased roles for non-animal alternatives, such as high-throughput robotic molecular and cellular screening methods.

It is expected that such in vitro methods will play a major role in future toxicity testing, in combination with SAR-based predictions, in silico biokinetic modeling, and newer, developing technologies such as genomics and systems biology (Blaauboer, 2008).

**Reduction alternatives**

Strategies designed to achieve reduction of laboratory animal use may be applied at three different levels, as we recently described in detail (De Boo and Hendriksen, 2005):

*Intra-experimental reduction* may be applied at the level of the experiment, directly. Reduction may often be achieved through improvements in experimental design and statistical analysis. Insufficient sample sizes presently leave many studies underpowered, limiting the statistical validity of study conclusions. Animal lives and other resources may also be wasted, if experiments subsequently require repetition. The relatively poor statistical knowledge of many animal researchers may be the cause of the high prevalence of poor sample size choices, within animal studies. Solutions could include the training of researchers in statistics, and the direct input of statisticians in experimental design and data analysis.

Strategies may also be considered for minimizing animal numbers without unacceptably compromising statistical power. Meta-analysis of multi-factorial and randomized block designs involves the aggregation and statistical analysis of suitable data from multiple experiments. For some purposes, treatment and control groups can be combined, permitting group numbers to be minimized.

Pre-screening of test compounds using replacement methods can result in a significant reduction in animal numbers. When information about the size and variability of the responses is already available, smaller, focused studies may be appropriate. Modifications should be made in routine tests to reduce animal use, for example, as a result of an increase in expertise, further standardization of procedures, and/or availability of animals with lower inherent variability. Analysis of pilot test data may yield information about data variance, which can then be used to calculate sample sizes.

Data variability, and hence the need for larger samples, may be decreased by minimizing heterogeneity within experimental environments and protocols. This can be achieved by: i) the appropriate use of environmental enrichment, aimed at decreasing physiological, psychological or behavioral variation resulting from barren laboratory housing and stressful procedures; ii) choosing, where possible, to measure variables with relatively low inherent variability; iii) the use of genetically homogeneous (isogenic or inbred) or specified pathogen-free animal strains; and iv) screening raw data for obvious errors or outliers.

*Supra-experimental reduction* involves changing the setting in which a series of experiments takes place. Examples include: improving the education and training of staff, such as training researchers in experimental design and statistics; including a ‘named statistician’ in animal ethics committees, achieving reductions in breeding surpluses, and critically analyzing, and re-defining, test specifications.

Policies encouraging the re-use or recycling of animals (so-called ‘4th R’) represent a special case of supra-experimental reduction. Several procedures allow for longitudinal studies, where animals may sometimes act as their own controls. This can make the use of parallel control groups, as well as interim kills, unnecessary. In such cases, however, the impact on recycled animals may be increased, and so a weighted case-by-case cost/benefit assessment must be made, to ensure that such impacts do not exceed the benefits achieved by minimizing numbers. It may, for example, result in higher animal welfare and ethical standards, when a lower impact on a larger number of animals is incurred, than a severe impact on a few.

At departmental level or even across departments within institutions, better planning of research activities, and cooperation and sharing of resources
between research programs, may further decrease animal use.

Extra-experimental reduction refers to reductions achieved through more distantly-related developments, such as improved research or production strategies that may result in superior quality, consistency and safety of pharmaceuticals, thereby decreasing data variability and test group sizes, as well as the perceived ‘need’ for animal testing.

The development of quality-controlled manufacturing processes for the production of biological products such as vaccines provides a good example. In this case Good Manufacturing Practice and Quality Assurance principles, which include testing at various checkpoints within the production chain, aim to ensure the consistency of each batch of products, thereby decreasing the safety and efficacy testing required for the final products (Hendriksen, 2006; Knight, 2008b).

Refinement alternatives
Experimental refinements include the use of analgesic and anesthetic modalities that will not unduly alter experimental outcomes of importance. Such modalities remain underutilized. Around 60% of procedures are conducted without anesthetics within the UK, for example (Hudson and Bhogal, 2006). While the use of anesthetics or analgesics undoubtedly alters normal physiology, claims that such alterations are sufficiently important to hypotheses under investigation, to warrant their exclusion, require careful scrutiny. It may, for example, be the case that physiological variations in response to painful stimuli are of greater consequence.

Other refinement strategies include: the use of non-invasive imaging modalities; the use of telemetric devices to obtain information remotely—although such devices may themselves compromise welfare, due to surgical trauma and the weight of the device within or on the animal’s body, so case-by-case assessments are necessary; fecal analysis (e.g. fecal cortisol may be analyzed, in some cases eliminating the need to blood-sample); positive reinforcement techniques such as training animals (especially primates) to participate in monitoring (e.g., presenting arms for venipuncture (blood-sampling)), rather than using physical or chemical restraint; environmental enrichment; and the opportunity to socialize when possible, with compatible conspecifics (others of the same species), for social species (De Boo et al., 2005).

Interaction between alternatives
When developing research questions and designs it is advisable to consider all alternative methods for specific procedures together, rather than in isolation. There are instances in which the use of one of the 3Rs can work in tandem with one or both the other two Rs (De Boo et al., 2005). For example, introducing education and training programs for staff is likely to improve animal care, assist with early detection of welfare problems, and aid in monitoring of experimental effects. This may result in decreased animal stress as well as improved recording of results, which may, in turn, reduce variability of results, and hence, animal numbers required.

The international harmonization of protocols and legally required safety tests represents an important strategy for reducing the number of animal experiments that must be carried out in individual countries. It is also a strategy by which obsolete, invasive animal studies can be replaced by the most advanced techniques using sentient material.

Similarly, the replacement of toxicity tests on live animals by animal tissue or cell cultures not only reduces animal numbers, but can also minimize the harm experienced by those animals that are used. Reduced costs and timeframes, and increased human predictivity, may sometimes result. In such cases all 3Rs may converge, for the benefit of science, industry, consumers and animals.

When implementation of one of the 3Rs would conflict with another potentially beneficial alternative method, researchers and ethical review committees are left to evaluate the impact of one method in preference to another. They may be required to make difficult decisions concerning the prioritization of alternative methods. For example, as a trade-off for replacing live animal use, fetal bovine serum is commonly used within in vitro assays, but the collection of blood may cause suffering to the fetus. The collection method could be modified, or a serum-free culture used, which offers additional scientific benefits (Evan et al., 2006). Another example of a conflict between reduction (through increased data acquisition from
fewer animals), and refinement, may arise from the use of telemetric devices, which, as mentioned previously, may occasionally produce welfare problems.

**Increasing the implementation of alternatives**

A variety of additional strategies could increase general compliance with 3Rs principles. In many cases particular alternatives have been developed and successfully used at certain laboratories, but transfer of such technology elsewhere is minimal. Reproducibility and transfer would be assisted by increased description, and standardization, of methodologies. Journal space is limited, and so Gruber and Hartung (2004) proposed the establishment of publicly-accessible methodology databases, which could provide supplementary links to articles—a technologically simple concept.

Public funds previously spent on animal tests required by regulators could also be redirected into the further development and implementation of alternatives (Gruber and Hartung, 2004).

As Gruber and Hartung observed, however, more fundamental problems appear to exist. These include incorrect assumptions of the human utility of animal models, and a dearth of interest in exploring alternatives to animal use. Considerably more stringent compliance with relevant animal welfare legislation requiring the consideration or use of alternatives could—and should—become a prerequisite of research funding, ethics committee approval, and publication of results. These measures would require the education and cooperation of funding agencies, ethics committees and journal editors, about the limitations of animal models, and the potential of alternatives (Knight, 2008b).

**Conclusions**

The scientific and logistical limitations incurred by the use of animal models of humans within biomedical research and toxicity testing are substantial, and increasingly recognized; as is social concern about, and consequent regulatory restriction of, laboratory animal use. In defiance of these factors, such use remains enormous. The use of GM animals, and the implementation of large-scale chemical testing programs, are increasing laboratory animal use internationally.

These trends clearly demonstrate the need for considerably greater awareness of, and compliance with, the principles of the 3Rs. These principles are universally recognized as essential to good laboratory animal practice, for animal welfare-related and ethical reasons, and also, to increase the quality of the research, and the robustness of subsequent results.

**References**


Rusche, B. (2003). The 3Rs and animal welfare—conflict or the way forward? *ALTEX* 20(Suppl. 1), 63–76.


**Corresponding authors:**
Jasmijn de Boo and Andrew Knight
Animal Consultants International
91 Vanbrugh Ct.
Wincott St.
London SE11 4NR
UK
Contact via:
www.AnimalConsultants.org