Chapter 9

Animal use in toxicity studies

Animal use in toxicity studies

Introduction

9.1 In this chapter we describe the purpose and principal methods of toxicity studies. Most of these studies are conducted to assess the degree to which substances are toxic (poisonous) for humans, animals or the environment, to investigate the mechanism of toxic chemicals, or to develop new or improved tests for specific types of chemically induced effects. We begin by explaining the scientific rationale behind important types of studies. These include: examination of adverse effects that may occur on first exposure to a single dose of a substance (acute toxicity studies), studies that seek to assess the potential of substances to interact with genetic material (genotoxicity), tests that aim to identify whether toxicity occurs after continuous exposure to a substance (repeated-dose toxicity studies), tests that are undertaken to find out whether cancers may develop as a result of exposure to certain chemicals, and studies to ensure the safety of medicines.

9.2 In the second part of the chapter we discuss a range of welfare implications that may arise for animals involved in toxicity testing. We consider first effects that may result from the dosing and sampling methods that are commonly used, and then effects related directly to the toxicity of the chemical that has been administered. Toxicity studies are highly variable in design, and where they involve the use of animals the implications for animal welfare must be considered on a case by case basis. We concentrate here on the more standardised animal methods that are widely used to characterise the adverse effects of chemicals on human and animal health, and on the environment. Many of the tests described are also used in the testing of medicines. For the most part we do not differentiate in the description between these different purposes.

Box 9.1: Toxicity studies – number of animals used

Procedures for toxicological purposes accounted for 16 percent of all animal procedures undertaken in 2003 in Great Britain. Approximately ten percent of all animal procedures were carried out for pharmacological safety and efficacy studies.

About 13 percent of all animal procedures were toxicological tests conducted to conform to legislative or regulatory requirements.

Source: Home Office (2004) *Statistics of Scientific Procedures on Living Animals Great Britain 2003* (London: HMSO).

The current approach

9.3 The vast majority of toxicity testing is carried out in the context of regulatory requirements governing particular types of chemical in different parts of the world (see paragraphs 13.49–13.51). Regulatory bodies often emphasise the necessity of toxicity tests to preserve current levels of human health and environmental protection.[1] Notification of all new chemicals placed on the EU market for the first time must be given to the competent authorities of the Member States.[2] The EU is currently considering a proposal for a Regulation concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) which would require large numbers of existing chemicals to be evaluated for safety. Reach would be binding for all Member States (Box 9.2). Separate European Directives specify requirements for animal testing in the authorisation or licensing of plant-protection products, biocides and pharmaceuticals (see paragraphs 13.49–13.51). In other

1 European Commission (2004) *Opinion of the Scientific Committee on Toxicity, Ecotoxicity and the Environment on The BUAV-European Coalition to End Animal Experiments Report: The Way Forward - Action to End Animal Toxicity Testing*, available at: http://europa.eu.int/comm/health/ph_risk/committees/sct/documents/out217_en.pdf. Accessed: 26 Apr 2005.

2 The information that must be supplied by a manufacturer is laid down in the Dangerous Substances Directive (67/548/EEC), implemented in the UK by the Notification of New Substances Regulations 1993.

cases, for example for cosmetics and cosmetic ingredients, testing requirements are not specified in regulations but there is a general requirement for safety, which could be met by the use of animal or non-animal tests. National authorities in the EU issue guidance on how the provisions laid out in the Directives should be met, which, due to preferences of regulators, usually means that data from established animal tests must be provided. So as to maximise returns, many chemicals are marketed worldwide, and testing must then conform to the requirements of other regulatory bodies, particularly those of the USA and Japan.

9.4 Current testing regimes have evolved significantly over the past three decades. Existing practices have changed and new methods have been added. A major influence on these developments has been the Test Guidelines Programme of the Organisation for Economic Cooperation and Development (OECD), which has developed standardised methods of testing that are accepted in principle by all 30 OECD Member Countries[3] through an agreement on the mutual acceptance of data.[4] The OECD approach has largely removed the need for testing according to different protocols to satisfy regulatory authorities in different countries, and has thus substantially reduced the total number of animals used for certain standard tests. It also provides a focus for the introduction of new methods that replace, reduce or refine animal use. Change and revision have been slow but there are many current initiatives, within both the scientific and regulatory communities, that challenge present practice with the aim of providing the same or even better levels of human safety while using fewer animals (see Box 2.4 and paragraph 11.10). The International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) also seeks to standardise the approach to testing of pharmaceuticals (see paragraphs 12.8, 13.50 and 15.84).

Box 9.2: The EU REACH Initiative: Registration, Evaluation and Authorisation of Chemicals

Registration, Evaluation and Authorisation of Chemicals (REACH) refers to the new EU regulatory framework for chemicals proposed by the EC in October 2003. At the time of writing, the proposal is being considered by the European Parliament and the Council of the EU. The legislation is intended to bring 30,000 chemicals manufactured within or imported into the EU under a single regulatory regime. REACH aims to make manufacturers responsible for the chemicals that they produce and to make it easier for highly toxic chemicals to be removed from the market. Under the new system, businesses that manufacture or import more than one tonne of a chemical substance each year would be required to register it in a central database.*

The European Commission has stated that new legislation is necessary due to the inadequacy of the current legislative framework for chemicals. A particular problem concerns the arbitrary cut-off date in 1981, which provides the distinction between 'new' and 'existing' chemicals. At present, 'new' chemicals that have been placed on the market after 1981 must be tested if their production exceeds 10 kg per year, whereas there are no such provisions for 'existing' chemicals. Therefore, it is argued, the current legislation encourages the continued use of untested existing chemicals because it is easier and cheaper.†

REACH has proved controversial, not least because its requirements will result in a substantial increase in the number of animal experiments. Many chemicals have been in use for decades and there is concern that some tests may duplicate those already performed by private companies. The UK Government is advocating a policy of 'one substance-one Registration', as a means of minimising animal testing and reducing costs and bureaucracy. This means that companies would be required by law to share data on tested substances, in the hope that universally available data will avoid duplicate testing of that substance.‡

* See EC Enterprise and Industry (2005) *The New EU Chemicals Legislation – REACH*, available at: http://europa.eu.int/comm/enterprise/reach/overview.htm.. Accessed on: 3 May 2005.

† EC (2003) *Q and A on the new chemicals policy REACH*, available at: http://europa.eu.int/rapid/pressReleasesAction.do?reference=MEMO/03/213&format=HTML&aged=0&language=EN&guiLanguage=en. Accessed on: 27 Apr 2005.

‡ See House of Commons Science and Technology Committee (2004) *Within REACH: The EU's New Chemicals Strategy* (London: TSO), available at: http://www.publications.parliament.uk/pa/cm200304/cmselect/cmsctech/172/172.pdf. Accessed on: 3 May 2005.

3 Member Countries include the UK and other European Countries, Japan and the USA, a list is available at: http://www.oecd.org/document/58/0,2340,en_2649_201185_1889402_1_1_1_1,00.html. Accessed on: 26 Apr 2005.

4 OECD (2001) *Mutual Acceptance of Data (MAD)*, available at http://www.oecd.org/document/41/0,2340,en_2649_201185_1890473_1_1_1_1,00.html. Accessed on: 26 Apr 2005.

9.5 Toxicity has two main components: the effect caused and the level of exposure (dose) at which the effect is observed. Some tests are designed specifically to detect a particular effect (such as skin and eye irritancy, skin sensitisation and mutagenicity studies). Other tests (such as sub-chronic and chronic studies) are designed to detect a wider range of less-specific effects on organs or body systems and the dose range over which the effect develops.

9.6 Information from toxicity tests is first used to provide a classification for a chemical, for example to assign appropriate warning labels for containers, and, where necessary, for selecting measures, such as protective equipment, during manufacture, exposure and use. Data from tests that characterise the relationship between dose and toxicological response are integrated with information on human exposure to produce a risk assessment, and to identify control measures necessary to manage and reduce any identified risk. Tests on species such as fish and amphibians are used in a similar way to assess the potential environmental effects of chemicals. For pharmaceuticals, results from animal tests are used in combination with data on the efficacy of a potential medicine to decide whether the beneficial effects of the treatment would outweigh the risks of adverse side effects, and to establish a safe dose for use in clinical trials (see paragraphs 8.26–8.28). They may also indicate potential side effects that must be monitored carefully.

9.7 The prediction of the likely effects of chemical exposure on human health is based primarily on the results of tests involving experimental animals. The number of animals involved in these tests varies. A full complement of toxicity tests for a successful pharmaceutical compound that proceeds to the market, involving single dosing, repeat sub-chronic and chronic dosing, reproductive testing, genotoxicity and carcinogenicity testing, can involve between 1,500 and 3,000 animals. The actual numbers required will depend on the need for further tests according to the nature of the test substance and also its toxic properties. The numbers of animals used to test other types of chemical are generally lower, but in some cases, where there is particular controversy about the safety of a chemical, tests may be repeated, with modifications, resulting in the use of even more animals.

9.8 Large numbers of animals are also used in several other tests. For example, a carcinogenicity bioassay generally involves 800 animals in total (400 of each sex) and may be conducted on both rats and mice. Adult animals (typically at least 80 animals of each sex per study), offspring and fetuses are used in reproductive and development studies. Rats and mice are most commonly used (74 percent), but in some cases testing is carried out on other animals such as rabbits (four percent), guinea pigs (three percent), dogs (one percent) or primates (less than one percent).[5] The interpretation of the results for assessing human safety depends on a number of assumptions. First, unless there is specific knowledge of species differences in the test response, it is assumed that the effects detected in rodents or other species are the same as those that would be induced in humans. Secondly, it is assumed that the sensitivity of the test animals represents, at best, the average sensitivity of the highly heterogeneous human population and that for some members of the human population the health risk could be much higher (see paragraph 8.39, Box 9.3 and paragraph 10.33). We consider next a range of examples to illustrate the different kinds of toxicity tests which are currently used.

5 Figures refer to percentages of procedures started in 2003 and carried out for the purpose of toxicology or safety and efficacy evaluation. See Home Office (2004) *Statistics of Scientific Procedures on Living Animals Great Britain 2003* (Norwich: HMSO).

Box 9.3: Sources of uncertainty in animal toxicity tests

Commentators who are critical about the reliability of toxicity tests carried out in animals with regard to predicting toxic effects in humans observe that the following factors may influence transferability:*

- species, strain and gender variations may affect extrapolation to humans;
- scaling from small, short-lived animals (usually rodents) and large doses, to large, long-lived animals (humans) and usually smaller doses may pose problems;
- there may be variability due to different dosing routes and extrapolation to human exposure;
- test animals usually constitute a homogeneous (genetic and otherwise) population, whereas there are often significant differences between humans, which may affect, for example, drug metabolism;
- there are pragmatic limitations with regard to testing chemical mixtures or interactions between chemicals; and
- the dose required to produce toxic effects in animals may never be reached in humans.

Researchers carrying out toxicological tests acknowledge that these factors need to be taken into account. They observe that, provided they are considered appropriately in making extrapolations, animals can be useful models for the prediction of toxic effects in humans (see paragraphs 10.27–10.36).†

* Langley G (2004) *Chemical Safety and Animal Testing: A regulatory smokescreen?* A BUAV Report (London: BUAV), available at: http://www.buav.org/pdf/Smokescreen.pdf. Accessed on: 4 May 2005.

† Association of the British Pharmaceutical Industry *(2001) Submission to the House of Lords Committee on Animals in Scientific Procedures*, available at: http://www.abpi.org.uk/information/industry_positions/011126.asp. Accessed on: 4 May 2005.

Principal types of animal-based toxicity tests

Acute toxicity

9.9 Acute toxicity refers to the adverse effects that occur on first exposure to a single dose of a substance. Separate tests are needed to detect the effects of contact with the skin and eye (corrosion, irritancy and sensitisation; topical or local toxicity) and the effects on internal organs of a substance that is swallowed, inhaled, absorbed through the skin or injected (systemic toxicity; see paragraph 9.28).

9.10 In the case of local toxicity, skin irritancy is normally assessed by applying the test substance to shaved areas of the backs of rabbits and observing the development of redness, swelling, erosion and ulceration over a period of 72 hours (see paragraph 9.36). Eye-irritancy tests involve administering the test substance directly into the eye of the rabbit and observing corneal opacity, swelling, reddening and other signs of irritation.

9.11 In the case of skin-sensitisation testing, multiple doses of the test substance are applied to the skin of guinea pigs to see if a later dose will cause a strong immune reaction, indicating sensitisation to the chemical (see paragraph 9.36). Tests using guinea pigs are increasingly being replaced by a test involving mice called the local lymph node assay. The test material is applied to the ears of the mice. After an interval the mice are euthanised and the early stages of sensitisation are detected by measuring the level of induced DNA synthesis in the lymph nodes. This test provides more useful information, uses fewer animals than the guinea pig test, and causes substantially less pain and distress to the animals involved.

9.12 The main purpose of skin and eye testing is to allow classification and labelling of corrosive, irritant and sensitising chemicals. The current systems of classification now follow a progressive, step-wise strategy that allows chemicals to be classified as corrosive or irritant to the skin by using physico-chemical properties, such as pH value. Tests that use isolated human or animal tissue cells or *ex vivo* tissues or organs to identify chemicals with the potential to cause severe irritation or corrosion are known as non-animal pre-screens (see paragraph 11.9). For sensitisation, analysis of chemical structure (structure–activity relationships) can identify many potential sensitisers. Therefore, in many cases it is now possible to classify chemicals without the need for animal tests. The value of these

approaches is illustrated by a decrease in rabbit eye tests in the UK from approximately 4,000 in 1995 to 1,100 in 2003.[6]

9.13 Acute systemic toxicity is assessed by the administration of a single dose of compound, typically to rats and mice, orally, dermally or by inhalation. For pharmaceuticals, the main aims of these studies are to determine the nature (including delayed toxicity) and duration of any acute toxic response. They also determine the maximum non-lethal dose and provide preliminary information relevant to single exposure or over-dosage in humans (see paragraph 9.39).[7]

9.14 For industrial chemicals and agrochemicals, testing covers acute toxicity by oral, dermal and inhalation routes of exposure. The information obtained is used primarily to ascribe a chemical to bands of acute toxic effect, which restricts how the materials may be used, and thus the extent of human exposure by the routes of exposure which have been evaluated. In the past, in the UK and elsewhere, acute systemic toxicity was investigated by the use of lethal-dose tests, in which the oral dose causing the death of 50 percent of the treated animals (the LD_{50} value) was determined.[8] Such tests used at least 30 animals per test chemical and required death of the animals as an endpoint, regardless of the suffering caused. In 2001 the OECD agreed that the LD_{50} test for acute oral toxicity should be abolished and deleted from the OECD manual of internationally accepted test guidelines by the end of 2002 (see paragraphs 9.4 and 12.8).[9] Several alternative methods have been developed which use fewer animals and in some cases replace death as the endpoint with signs of significant toxicity instead. Information on similar chemicals is used to guide the selection of initial dose levels and the tests are designed to avoid or minimise lethality or severe toxicity. These methods have replaced the LD_{50} test for acute oral toxicity, but several acute tests such as those involving inhalation, dermal and eye exposure have yet to be modified. They are still used internationally for tests on birds and, for some purposes, also on mammals.[10] Lethal-dose tests are also still used to assess the safety of biological products, such as vaccines (Box 8.5), and certain foods, such as shellfish, for the presence of toxins (see paragraph 9.37).

9.15 The approach to assessing the acute toxicity of pharmaceuticals differs from that described above, in that maximum tolerated dose (MTD) studies are carried out to aid the later process of dose selection. These tests often replace acute studies, especially in the case of larger species such as the dog and primates which are used to complement and verify earlier findings in rodents. They involve steadily increasing the dose given to an animal (single or a number of consecutive doses), until adverse effects indicate that an MTD has been reached. This is normally determined by careful observation of the animals, but there is no universally accepted definition of the MTD and effects such as vomiting and convulsions may occur and are sometimes used as signs of the MTD (see paragraphs 9.34–9.45).

---

6 See Home Office (2004) *Statistics of Scientific Procedures on Living Animals Great Britain 2003* (Norwich: HMSO).

7 The studies provide information that may support selection of dose levels for repeated-dose toxicity studies, *in vivo* genotoxicity tests (see paragraphs 9.20-9.21) and, subsequently, first human exposure studies.

8 The OECD gives the definition as the dose that can be expected to cause death in 50 percent of animals when administered by the *oral* route.

9 OECD (2001) *OECD Test Guideline 401 will be deleted: A Major Step in Animal Welfare: OECD Reaches Agreement on the Abolishment of the LD_{50} Acute Toxicity Test*, available at: http://www.oecd.org/document/52/0,2340,en_2649_34377_2752116_1_1_1_1,00.html. Accessed on: 27 Apr 2005; The UK ceased the practice of the LD_{50} test in 1999 (APC (1999) *Press release: LD 50 Test - Changes To Licensing Procedures*, available at: http://www.apc.gov.uk/press_releases/991021.htm. Accessed on: 27 Apr 2005). In the US, Environmental Protection Agency guidelines describe the LD_{50} test as standard for pesticides and toxic substances although state that it may be unnecessary in certain circumstances (Environment Protection Agency (1998) *Health Effects Test Guidelines OPPTS 870.1100 Acute Oral Toxicity*, available at: http://www.epa.gov/docs/OPPTS_Harmonized/870_Health_Effects_Test_Guidelines/Series/870-1100.pdf. Accessed on: 27 Apr 2005.

10 OECD (2002) OECD guidelines for the testing of chemicals: Proposal for a new guideline 223 – Avian acute oral toxicity test, available at: http://www.oecd.org/dataoecd/16/41/1836204.pdf. Accessed on: 27 Apr 2005.

Repeated-dose toxicity studies

9.16 These studies have three main objectives (i) to identify toxicity that develops only after a certain length of continuous exposure to the chemical, (ii) to identify the organs most affected and (iii) to determine the doses at which each effect occurs.

9.17 Repeated-dose studies are conducted for various periods of time. The 28-day (sub-acute) study is most common, but studies of 90 days to one year are also regularly carried out. Rats and mice are generally used but for certain classes of chemicals, such as agrochemicals and pharmaceuticals, the tests may also be conducted in non-rodent animals such as the beagle dogs, pigs, marmosets or macaques (see paragraphs 9.26 and 9.30). The test data allow an assessment of the highest dose without significant effects (the 'no observed adverse effect level', or NOAEL). This is used in risk assessment and risk management, by limiting the acceptable exposure of humans to a fraction of the NOAEL. For example, in the case of agrochemicals and food additives, these studies are used to assign a reference dose to which safety factors are applied to give an acceptable daily intake (ADI) that is typically a hundredfold less than the observed NOAEL. This can be defined as the dose level to which humans may be exposed, through residues on foodstuffs and in drinking water, with the practical certainty that no adverse health effects will ensue.

9.18 Repeated-dose studies are also used to give an insight into any species differences in toxicity that could be relevant to the assessment of risk in human health. Depending on the use and physico-chemical properties of a chemical, different routes of administration, such as oral, by inhalation or dermal contact, may be used to give a more appropriate risk assessment. For pharmaceuticals, results of these studies support investigations requiring the first administration of the test substance to humans.

Carcinogenicity

9.19 For the assessment of carcinogenicity, rats and mice are dosed for up to two years (the typical lifespan for these species) and the incidence and type of the tumours that develop is evaluated (see paragraph 9.33). This knowledge is used to assess the risk of cancer induction by the chemical in exposed humans. In practice, the assessment of repeated-dose studies and carcinogenicity is often combined into a single study in rodents thus reducing the use of experimental animals.

Genotoxicity

9.20 Short-term studies investigating interactions with genetic material (DNA and chromosomes) are widely used to screen chemicals for the potential to cause cancer or heritable mutations. Most of these studies involve the use of *in vitro* assays for mutation in bacteria or isolated mammalian cells that have been shown to predict the potential for a substance to be carcinogenic or mutagenic through interaction with DNA. In the pharmaceutical industry tests are performed as high-throughput screens (see paragraphs 8.9 and 8.21), both early in drug discovery and also to support drug registration. Animal studies, usually in the mouse, are used only when one or more of these *in vitro* tests has given a positive result, and with the purpose of demonstrating that the chemical can or cannot reach a sensitive tissue and cause genetic changes in the intact animal. In practice, very few chemicals that have been confirmed to be mutagenic *in vitro* are tested any further in animals. However, in the case of pharmaceuticals, regulatory requirements demand that an *in vivo* test be completed before the start of Phase II clinical studies in humans.

9.21 *In vivo* tests include the rodent bone marrow micronucleus test, which is an early predictor of carcinogenic activity. A single dose of compound is administered to rats or mice which are killed either 24 or 48 hours later for examination of chromosomal changes in bone marrow

cells. It is expected that the highest dose level used will show evidence of adverse effects if the substance is genotoxic, and the MTD is normally used to set this dose level.

Effects on reproduction and development

9.22 Studies within this category are intended to determine the effects of compounds on various aspects of the reproductive capacity of the adult, and on the development of the offspring. The most comprehensive test method for reproduction (the two-generation reproduction study) involves repeated oral doses to young rats through the period of sexual maturation into young adulthood when the animals are mated to treated females. The females are dosed throughout pregnancy and until the offspring are weaned. The pups are dosed until adulthood and mated and the second-generation young evaluated. These tests provide information on fertility, mating behaviour, parental behaviour and development of the neonate to adulthood. The results are used in hazard classification and in risk assessment. More-limited information on fertility and reproductive performance can also be obtained from a one-generation study or a screening test, which combines reproductive investigations with a 28-day repeated-dose toxicity test.

9.23 Studies on developmental toxicity provide specific information on the potential hazards to the unborn that may arise from exposure of the mother to a particular substance during pregnancy. Typically, groups of pregnant rats or rabbits are treated orally for up to the whole period of gestation, and the uterine contents are then examined and evaluated just prior to parturition. An evaluation is made of maternal toxicity relative to that in non-pregnant females, embryo or fetal death, altered growth and structural changes in the fetus. Rabbits are used in addition to rats and mice because rodents do not generally respond, or respond variably, to the effects of potent human teratogens such as thalidomide (see Box 8.4). The results of both tests are used in classification by hazard and in risk assessment.

Safety pharmacology

9.24 In the development of pharmaceutical products, additional tests are needed to detect exaggerated intended or unintended pharmacological responses. Pharmacology studies to evaluate safety are generally conducted in the dog (for cardiovascular endpoints) and in the rodent (for assessment of the effect on the whole body). Some examples of the types of studies performed in animals are described briefly below (see Box 9.4).

- **Dog telemetry:** dogs are implanted with radio-transmitters for continuous monitoring of blood pressure, heart rate, body temperature and electrocardiogram (ECG, see paragraph 4.56). These parameters can be monitored from the conscious dog, and allow the measurement of the effects of test compounds on the cardiovascular system *in vivo*. Dogs are usually reused in multiple studies, subject to veterinary and regulatory approval by the Home Office. They are euthanised at the end of the studies.
- **Haemodynamics of anaesthetised dogs:** this type of research is undertaken as a follow-up to dog telemetry. Under terminal anaesthesia, multiple systems may be investigated including, for example, blood pressure, heart rate, ECG and peripheral blood flows (also coronary and renal blood flows).
- **Absorption, distribution, metabolism and excretion studies (ADME):** although not strictly toxicity studies, these investigations (typically undertaken in rodents and dogs) are used to assess the amount of chemical or pharmaceutical that is absorbed into the animal, where it is distributed within the body, how it is changed by metabolism, the time-course for these events and how, and at what rate, the material is eliminated from the body (see paragraph 9.31). This information is used to select dose levels for toxicity studies and

clinical trials, to identify compounds for further development, to interpret toxicity data, and in risk assessment (see paragraphs 8.10–8.11).

- **'Balance' studies:** in these studies radiolabelled doses are given to intact or surgically prepared animals and samples including blood, bile, urine, faeces and expired air are collected to determine the processing of the drug-related material, and to investigate its absorption and possible retention.
- **Pharmacokinetic studies:** studies are conducted for pharmacological and toxicological evaluation of candidate drugs to characterise their pharmacokinetic behaviour, usually after intravenous and oral administration, although other routes may also be used (see paragraphs 8.20–8.26). This information is used to support the more limited sampling performed in toxicity studies, to fully characterise the pharmacokinetics in animals and to predict the pharmacokinetics in humans, which assists in estimating the likely human dose.

Box 9.4: Example of research – testing species differences in the toxicity profile of an approved herbicide (currently in use)

Lappin GJ, Hardwick TD, Stow R *et al*. (2002) Absorption, metabolism and excretion of 4-chloro-2-methylphenoxyacetic acid (MCPA) in rat and dog *Xenobiotica* **32**(2): 153–63.*

This research investigated differences between rats and dogs in the toxicity of a herbicide, MCPA. This chemical is used to control a wide variety of broad-leaved weeds in many crops as well as non-crop areas. A radioactive version of the herbicide was fed to the rats and dogs.

Twenty rats between six to eight weeks old and four beagle dogs between six and 12 months of age were used, obtained from suppliers of laboratory animals in the UK.

Two groups of rats were administered single doses of the herbicide at different levels by gavage (feeding by means of a stomach tube, see paragraph 9.28). Half the rats were group-housed in the period following dosing and a sample of their blood was taken on ten occasions. The remaining rats were housed individually, and their urine and faeces were collected for seven days.

For all four dogs a single dose was administered by capsule, followed by a second single dose at a higher concentration four weeks later. All four dogs were housed individually for five days following dosing, during which time their blood was sampled at 11 time points, and samples of urine and faeces were collected.

Signs of toxicological response to this compound had previously been shown to include reduced weight gain, increased kidney weight and altered clinical chemistry in the rat. The effects in the dog were more severe with clear hepatotoxicity (having a damaging effect on the liver), anaemia and severe renal toxicity. The highest dose given in this procedure resulted in mild toxicological effects in the rats. The responses in dogs were described as being beyond the MTD if repeated exposures at this level had occurred.

The researchers found that MCPA did not accumulate in rat tissue. The results were less clear in the case of the dog as this species is more sensitive to the effects of MCPA. The authors reached the most probable physiological explanation for the species differences. They also investigated previous evidence that this type of compound may reach higher blood concentrations in males than females, and found that there were in fact no differences. For this reason the researchers went on to use only male dogs, rather than increase the number of dogs used. The authors state that the data add to a growing body of evidence showing that the dog is deficient in the excretion of weak organic acids, and that therefore this species is not appropriate for assessing the toxicological significance of this class of compound in humans.

* This is an example of animal research that has been carried out in the UK and published in a peer-reviewed journal. Details relate to this specific example and should not be taken to represent a 'typical' animal experiment. It is important to note that individually published experiments usually form one part of a continuing area of research, and the significance of the results may therefore be difficult to interpret.

Ecotoxicity

9.25 All of the tests described above are carried out to assess the possible adverse effects of a substance on human health, but an increasing amount of testing is being done to investigate potential effects on the environment and wildlife. For example, large numbers of fish, and smaller numbers of birds and amphibians, are used to test industrial and agrochemicals for their toxicity to wildlife populations (see also Box 9.4).

Issues concerning the welfare of laboratory animals in toxicity testing

9.26 We have commented on the numbers and types of animals most commonly used in toxicity testing (paragraph 9.8). We also observed that some toxicity tests may extend over several months or years in contrast to most animal experiments conducted for biomedical research. For rodents, age-dependent health problems, with concomitant stress, will usually occur with increased frequency towards the end of tests. Loss of animals can compromise study validity and confound the interpretation of data, especially from carcinogenicity studies.[11] This may sometimes encourage investigators to minimise animal loss by avoiding euthanasia as far as possible, which may result in increased pain and distress to the animals.

9.27 It is impossible to fully predict the pain and suffering that individual animals might experience during toxicity testing. However it is possible to assess the likelihood that pain and distress will occur under a particular set of conditions and exposures. The following aspects of toxicity testing can give rise to adverse consequences for the welfare of test animals, the extent of which depends on the test and species involved: (i) transport (see paragraph 4.36); (ii) housing and husbandry (see paragraphs 4.37–4.43);[12] (iii) dosing and sampling procedures (which might be repeated) (see paragraphs 4.49–4.52); (iv) the length of the observation period and (v) the toxic consequences of dosing. The adverse effects on animals that may arise specifically in toxicity tests, as opposed to other forms of animal research, are due mainly to dosing procedures and the toxic effects of the treatments.[13]

9.28 Dosing can involve the repeated administration of test material by a variety of routes of exposure, including gavaging (stomach intubation or forced feeding), injection, skin painting and inhalation. Some types of administration are likely to be very stressful to animals, especially when they are repeated and are of relatively long duration (see paragraphs 4.45 and 9.28). In addition, dosing into the eye and inhalation exposure involve restraint for several minutes or hours.

9.29 The right choice of dosing vehicle and volume is an important means of refining toxicity tests from both scientific and welfare perspectives. This is particularly so regarding the maximum amounts that should be administered to the eye and orally by gavage.[14] The use of low dosing volumes is a very effective way of reducing stress during topical ocular administration. Thus, the traditional dosing volume of 0.1 ml can be reduced by a factor of 10 or even 20 in eye-irritation studies. During gavaging, volumes of 1–50 ml/kg are usually administered, depending on the species being used. The administration of large volumes through this route can modulate the patterns of absorption, thereby affecting toxicity. For example, volumes nearing or exceeding the stomach volume will result in the delivery of some of the substance to the small intestine.

9.30 Stress can also be induced by physiological changes accompanying oral dosing. For example, alterations to gastric secretion and motility, as well as increases in heart rate and blood pressure, can occur. There can also be changes in biochemical parameters, such as levels of

11 Roe FJC (1993) Influence of animal species, strain, age, hormonal, and nutritional status, in *Experimental Toxicology, The Basic Issues*, 2nd Edition, Anderson D and Conning D (Editors) (Cambridge: The Royal Society of Chemistry), pp23–34.

12 Morris T, Goulet S and Morton D (2002) The international symposium on regulatory testing and animal welfare: recommendations on best scientific practices for animal care in regulatory toxicology *ILAR J* **43**, Supplement: S123–5; Hawkins P, Morton DB, Bevan R *et al.* (2004) Husbandry requirement for rats, mice, dogs and non-human primates used in telemetry procedures Seventh Report of the BVAAWF/FRAME/RSPCA/UFAW Joint Working Group on Refinement, Part B. *Lab Anim* **38**: 1–10.

13 Stephens ML, Conlee K, Alvino G and Rowan AN (2002) Possibilities for refinement and reduction: future improvements within regulatory testing *ILAR J* **43**, Supplement: S74–9.

14 Brown AP and Levine BS (1999) *Relationship Between Dosing Vehicles, Dose Volume, and Stress*. Report prepared for the US National Toxicology Program Unpublished report.

stress hormones. Furthermore, under conditions where animals are fed in laboratories *ad libitum*, as is the usual situation, gavaging of large volumes may result in aspiration of the test substance due to the presence of food in the stomach and duodenum. The volume of the gastrointestinal tract for receiving administered material is reduced and injury to the lungs may ensue. Recent research showed that gavaging rats with corn oil, but not the test substance or water, resulted in stress which was volume-dependent, as manifested by corticosterone levels (a hormone released in response to stress).[15] The authors recommended that dosing volumes for rats should not exceed 10 ml/kg. It is important to consider this information in the light of other best-practice guidelines on dosing.[16] At the same time, views differ as to how widespread the gavaging of large volumes *ad libitum* is in practice, and some researchers comment that significant steps have been made to refine the method.[17]

9.31 In metabolism studies, animals are housed in metabolism cages and might have external tubes implanted into their bile ducts.[18] During toxicokinetic studies in dogs, it is not unusual for the same animals to be reused after a suitable period of time, as such animals are thought to suffer less stress than those used for the first time.

Effects due to toxicity

9.32 The usual practice in toxicity testing is to induce overt toxicity in some animal groups, in order to ensure that, where toxicity is not observed in other exposed groups, the effects are not due to any inherent defect in the methodology. Thus, some form of harm to animals is an integral part of animal-based toxicity testing and is viewed by those conducting such tests as being unavoidable to achieve the scientific objectives of the work.

9.33 Toxicity can arise from reversible or irreversible effects, and can affect a range of different organs to different degrees. The adverse effects of substances on animal physiology can range from minor changes, such as reduced weight gain, small physiological alterations or changes in the levels of circulating hormones, to severe effects such as organ function loss (a major cause of acute toxicity), leading to death. Intermediate levels of toxicity, such as those destroying tissue and adversely affecting tissue function, could result in pain and suffering. Similarly, the development of tumours during carcinogenicity testing, or intestinal swelling during sub-chronic or chronic testing, might also lead to pain and discomfort.

9.34 The adverse effects which are used to define the MTD range from the very mild, which include non-clinical signs of lethargy or effects on weight, to the more substantial, such as convulsions. For example, various tests of toxicity often require signs to be scored, such as changes in the condition of the coat and eyes, as well as other signs of ill-health. Many of these conditions might be expected to reflect pain and suffering to differing degrees.

9.35 There is general confusion among toxicologists as to exactly what defines an MTD, 'severe distress', 'obvious pain', a 'moribund condition' and other descriptions of animal welfare. Some have argued that the relevant OECD test guidelines need to be revised accordingly.[19] Several of the OECD test guidelines are vague on issues such as environmental enrichment, where for example group housing is not specified when it would be possible.[20] All these

15 Brown & Levine (1999) ibid.

16 Morton DB, Jennings M, Buckwell A *et al*. (2001) Refining procedures for the administration of substances *Lab Anim* **35**: 1–41.

17 See, for example: Brown AP, Dinger N, Levine BS (2000) Stress produced by gavage administration in the rat *Contemp Top Lab Anim Sci* **39**:17-21.

18 Gangolli SD and Phillips JC (1993) The metabolism and disposition of xenobiotics, in *Experimental Toxicology, The Basic Issues*, 2nd edn, Anderson D and Conning D (Editors) (Cambridge: The Royal Society of Chemistry), pp130–201.

19 Koeter HBWM (1999) The OECD Test Guidelines Programme and animal welfare concern: how to avoid major animal suffering, in *Humane Endpoints in Animal Experiments for Biomedical Research*, Hendriksen CFM and Morton DB (Editors) (London: Royal Society of Medicine Press), pp13–14.

ambiguities can act as potential sources of avoidable suffering for the animals.

9.36 Other examples of toxicity endpoints that are likely to be painful and stressful include skin irritation and corrosion where single doses are applied to shaved areas of the backs of rabbits. Exposure can extend over four hours, and the animals may experience ulceration of the skin as well as swelling and itching. In sensitisation testing, multiple dosing is practised, and in addition to the above signs, the skin may crack and peel. Other signs that can be observed during acute, sub-acute and chronic toxicity testing include both external and internal bleeding, diarrhoea, loss of appetite, vomiting (in non-rodents), aggression, salivation, changes in blood pressure, coma, convulsions, lateral recumbency and tremors, loss of fur and hair, dehydration, or nasal discharge. Some of the less drastic effects of toxicity can arise merely from the act of dosing.

9.37 Very severe adverse effects can become manifest extremely rapidly as a result of neurotoxicity following dosing. For example, during the mouse bioassay for diarrhoetic shellfish toxins, atypical results[21] can arise which cause rapid death, following signs of substantial distress from shock and extensive trauma, accompanied by violent and rapid leg and body movements and agonal breathing (abnormal and uncertain respiration often characterised by gasping for breath), collapse and finally death from heart failure.[22]

General observations concerning the assessment of animal welfare in toxicity studies

9.38 It is difficult to assess accurately either the individual or the collective burden of suffering that is sustained by animals used in toxicity testing. Many toxicity procedures do not usually result in more than some discomfort to most of the animals concerned, at least in the case of rodents. Moreover, only certain test groups of animals will be subjected to tests leading to overt signs of toxicity during an experiment. These groups of animals comprise the concurrent positive controls (animals treated with a chemical known to have adverse effects as a comparator on the sensitivity of the test substance) and those animals that receive high doses in dose-response studies. However, in such cases it is likely that significant pain and distress could result, depending on the type of toxicity elicited. All animals used in toxicity testing are routinely killed immediately at the end of experiments for examination (see paragraphs 3.47–3.49).

9.39 The fact that animals can suffer stress during toxicity testing has been investigated in studies in rats by assessing stress and discomfort from clinical and pathological observations.[23] A substantial proportion of the animals suffered from serious discomfort, with some having obvious clinical signs, such as impaired locomotion and anaemia. Most of these animals only displayed non-specific clinical signs and the development of humane endpoints was confounded. The difficulty of interpreting data where overt toxicity is induced can be exacerbated by the fact that dosing of very high levels of test material might be required, with accompanying adverse welfare consequences for animals, including death. Death as an endpoint in toxicity testing, particularly when caused by the above conditions (the administration of 'heroic' doses), can be a misleading indication of hazard, since it might well not reflect any direct biological effects of the test material. Rather, death in such

20 Combes RD, Gaunt I and Balls M (2004) A scientific and animal welfare assessment of the OECD health effects test guidelines for the safety testing of chemicals under the European Union REACH system. *ATLA* **32**: 163-208.

21 These effects are 'atypical' in the sense that they arise very rapidly, usually within minutes of administration of the toxin (in most other cases effects more commonly occur within a timespan of several hours).

22 Combes RD (2003) The mouse bioassay for diarrhetic shellfish poisoning: a gross misuse of laboratory animals and of scientific methodology *Alternat Lab Anim* **31**: 595–610.

23 Van Vlissingen JMF, Kuijpers MHM, van Oostrrum ECM *et al*. (1999) Retrospective evaluation of clinical signs, pathology and related discomfort in chronic studies, in *Humane Endpoints in Animal Experiments for Biomedical Research*, Hendriksen CFM and Morton DB (Editors), (London: Royal Society of Medicine Press), pp89–94.

circumstances can be due to indirect effects such as dehydration leading to a heart attack. Similar effects can be caused by starvation which might occur when food becomes unpalatable during dietary administration of the test substance.

9.40 It has also been stressed that the design of toxicity experiments should be related to the way in which the resulting experimental data are going to be used.[24] Thus, if it is intended to label a substance as hazardous on the basis of adverse reactions detected in one or a few animals, there is little point in subjecting additional animals to treatment and potential toxicity. The use of pilot studies in which the unknown effects of a treatment can be assessed in a few animals prior to conducting a full-scale experiment are also desirable in order to reduce numbers of animals used, and the potential suffering. This approach is, unfortunately, not routinely practised by toxicologists.

9.41 It is important that those who care for and subject animals to toxicity testing should become aware of the behavioural, emotional and physiological conditions and requirements of the animals (see paragraph 4.18). The ability of animals to anticipate negative events such as experimental procedures can increase anxiety levels and alter hormonal production which might also compromise the scientific quality of the data.

9.42 Several factors are expected to increase the numbers of animals being used in toxicity testing, as well as the severity of testing, including:

- the High Production Volume chemicals testing programme in the USA;[25]
- the new Registration, Evaluation and Authorisation of Chemicals (REACH) legislation in the EU[26] (see Box 9.2);
- pesticide regulations in the EU that require more-extensive testing;
- the development and attempted validation of several animal tests to screen chemicals for endocrine (hormone)-disrupting activity;[27] and
- the very substantial increase in the generation and utilisation of novel GM animal strains in toxicity studies.[28]

Effective implementation of the Three Rs in these areas is crucial (see Chapters 11 and 12).

9.43 Finally, it must be acknowledged that toxicity tests in laboratory animals have limitations as a means of identifying hazards for human health, and managing risks to human health (see also Box 9.3). The example given in Box 9.4 also shows that different species may respond differently to the same compound. It has been argued that such problems fundamentally undermine the scientific and ethical justification for using animals to assess chemical safety. We have considered these questions briefly in paragraphs 8.39–8.41 and return to issues raised by the scientific validity of using animals in Chapter 10.

---

24 Morton DB (2002) The importance of non-statistical experimental design in refining animal experiments for scientists, IACUCs, and other ethical review panels, in *Applied Ethics in Animal Research: Philosophy, regulation, and laboratory applications*, Gluck JP, DiPasquale A and Orlans FB (Editors) (West Lafayette, IN: Purdue University Press), pp149–78.

25 Nicholson A, Sandler J and Siedle T (2004) An evaluation of the US High Production Volume (HPV) chemical-testing programme. A study in (ir)relevance, redundancy and retro thinking *ATLA* **32** Supplement 1: 335–41.

26 Combes R, Dandrea J and Balls M (2003) A critical assessment of the European Commission's proposals for the risk assessment and registration of chemical substances in the European Union *ATLA* **31**: 353–64.

27 Combes RD and Balls M (2003) How much flexibility is possible when validating new *in vivo* and *in vitro* toxicity test methods? *Alternat Lab Anim Exp* **31**: 225–32.

28 van Zeller A-M and Combes RD (1999) Transgenic mouse bioassays for carcinogenicity testing: a step in the right direction? *Alternat Lab Anim Exp* **27**, Supplement 1: 839–46.

Summary

9.44 In this chapter we have surveyed the ways in which animals are used in safety assessments of compounds including medicines, household chemicals, agrochemicals and industrial chemicals. Various species are used, most commonly rodents and also larger animals including rabbits, dogs and primates. Chemicals (including potential medicines) are assessed for their potential to be hazardous to humans, and estimates of the risk of adverse effects from particular levels of exposure are produced. Most toxicity testing is undertaken in the context of legal and regulatory requirements governing the use of particular types of chemical in different parts of the world.

9.45 A range of tests are described including: inhalation, skin irritancy, genotoxicity, acute dosing, repeated dosing and effects on developing fetuses. We observed that a full complement of toxicity tests for a pharmaceutical compound that reaches the market usually involves between 1,500 and 3,000 animals. Adverse welfare effects may arise from the environment in which animals are kept, and may therefore depend on housing and handling conditions (see Paragraphs 4.37–4.47). Specific welfare implications resulting from toxicity procedures depend on dosing and sampling methods, and the effects of the chemical. While toxicologists emphasise that many procedures affect animals only in minor ways, certain groups of animals, especially those in the positive control group, will be subjected to tests leading to overt signs of toxicity during an experiment, which means that significant pain and distress could occur, depending on the type of toxicity elicited. We consider ways of replacing, refining and reducing these effects in Chapters 11 and 12. In the next chapter, we summarise the discussion presented in Chapters 5–9, and consider in particular arguments about the scientific validity of animal research.