The use of animals in toxicity testing:
An RSPCA information paper

RSPCA policy

The RSPCA is opposed to all experiments or procedures that cause pain, suffering, distress or lasting harm to animals. This includes toxicity testing.

The Society's goal is to see all toxicity tests on animals replaced by alternative, non-animal methods. This brief report explains the RSPCA's approach to this issue – challenging the justification for subjecting animals to pain, suffering or distress in toxicity tests, and emphasising the need to develop new approaches to safety assessment without using animals.

Summary of RSPCA views on the use of animals in toxicity testing

- The RSPCA believes that the use of non-human animals to identify and quantify the human health risks of chemicals is a flawed approach. Evidence concerning the scientific validity of animal tests is fragmentary and insufficient to draw general conclusions, but toxicologists themselves acknowledge the difficulties associated with extrapolating from one species to another, and from carefully controlled laboratory experiments to real life situations.

- Despite their scientific shortcomings, animal tests currently represent the mainstay of most toxicity testing strategies. It is often claimed that their removal would seriously endanger human health. However, it is essential that new methods, not involving the use of animals, are developed as a matter of urgency to solve both the scientific and ethical problems associated with the use of animals. The RSPCA will do everything it can to promote the development and use of alternative methods.

- The results of animal tests done to satisfy legal requirements do not always prove to have been necessary or valuable, and are sometimes not acted upon when the safety of a substance is assessed, or safety measures decided upon. Animal tests should certainly not be done if it is uncertain that the information they produce will be of value, and will lead to effective action to minimise risks to humans, other animals and the environment.

- Some substances are tested on animals even though they have little or no real value to society. The suffering caused to animals in such cases cannot be justified.

- To be realistic, the phasing out of animal toxicity tests is likely to take a long time. In the meantime, millions of animals will spend their lives in laboratories, and will be subjected to painful and distressing procedures. The RSPCA will pursue all opportunities to reduce the numbers of animals used in toxicity testing and to alleviate the suffering of those who are used.

What toxicity testing means for animals

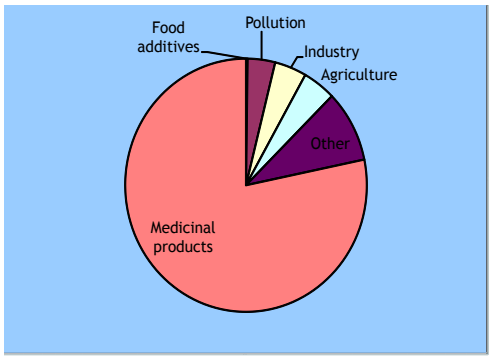
Statistics published by the Home Office[1] show that in Great Britain, 429,665 animals were used in 2009 in procedures described as "*toxicology and other safety/efficacy evaluation*". Many species were used including mice, rats, dogs, monkeys, birds and fish.

Animals used in toxicology in Great Britain, 2009	
Mice	188,083
Rats	117,040
Guinea pigs	6,360
Rabbits	7,325
Dogs	3,755
Non-human primates	2,122
Farm animals	3,431
Birds	13,694
Fish	85,566
Other species	2,289
Total	429,665

A large number of different tests are classified as toxicology procedures, but all are intended to find out what effects a chemical or product will have on the health of living animals. This is usually done to help predict the possible effects on people or the environment, including the animals in it.

Most (78%) toxicology tests are done on new medicines, or as a quality control measure for products such as vaccines. A substantial number of animals are also used in basic research on toxicity, and in the safety testing of chemicals used in industry and agriculture. Very few tests are done on household products, and none on cosmetics in the UK, although animals are still used in cosmetics testing in other countries.

Toxicity testing as a whole results in a great deal of animal pain, suffering, distress and death. There are many potential sources of stress and suffering within a laboratory environment, such as transport, confinement in laboratory cages, and even standard husbandry procedures such as cage cleaning. However, the biggest problem is the tests themselves.



Animals suffer pain and distress from the effects of being dosed with chemicals on their skin or into their eyes, by injection, by stomach tube, or by inhalation. For inhalation studies, rats or mice are often confined to small tubes and forced to breathe contaminated air for a number of hours.

Much more severe suffering will result if the substance being tested is poisonous. Some of the symptoms observed in toxicity testing include internal bleeding, diarrhoea, loss of appetite, aggression, salivation, changes in blood pressure, coma, convulsions, tremors, loss of fur and hair, dehydration, or nasal discharge. Very severe adverse effects can occur extremely rapidly as a result of nerve toxicity. For example, in a test for toxins in shellfish, mice may show signs of substantial distress from shock and extensive trauma, accompanied by violent and rapid leg and body movements, gasping for breath, collapse, and finally death from heart failure. Fortunately, few of the animals used in toxicity testing suffer to such an extent, but almost all animals are killed at the end of testing, for post mortem examination.

[1] Home Office (2010) Statistics of Scientific Procedures on Living Animals Great Britain 2009 (London: The Stationery Office)

Toxicity testing: Does the end justify the means?

The purpose of toxicity testing is to provide information on the biological effects of substances, so that precautions can be taken to protect humans, animals and the environment from the adverse effects of products used in medicine, industry, agriculture, and the household[2]. Some tests involving the use of live animals are intended to predict the likely effects of chemicals on species other than humans (environmental toxicology), but the majority are intended to provide information considered necessary for the protection of human health.

The main justification given for the use of animals in toxicity testing is therefore that the tests are necessary for the protection of human health, and that without animal testing the health of people exposed to chemicals during use, manufacture or transport would be at risk. The claimed justification for causing laboratory animals pain, suffering, distress and lasting harm in toxicity tests therefore rests upon a number of assumptions:

1. that the results of animal toxicity tests can be used with 'reasonable' reliability to predict the risk of adverse effects resulting from the exposure of humans to specific substances;
2. that the results obtained from animal tests will actually be used to decide on effective measures to protect human health, and that these measures will be implemented;
3. that the substance being tested is of sufficient value to society to justify the harms caused to the animals used to assess its safety;
4. that information of equivalent value in the protection of human health cannot be obtained by other means, not involving the use of animals.

The RSPCA believes that **all of these assumptions should be challenged**. We will take a closer look at each one in the following sections.

Assumption 1: Are animal tests scientifically reliable and valid?

There is an ongoing debate about whether toxicity tests on animals are, or are not, scientifically valid. Neither of the absolute positions on this question are sustainable or really helpful, a point made by both the Animals Procedures Committee (APC)[3] and the Nuffield Council of Bioethics working party on the Ethics of Animal Experimentation[4]. The RSPCA is, in any case, opposed to these tests on animal welfare and ethical grounds and wishes to see them replaced, *whether they are scientifically valid or not*. Nevertheless, the Society believes that it is essential to critically question the scientific validity of individual animal tests and their various applications on a case by case basis, since this will encourage and facilitate the development and use of alternative methods.

Challenging the scientific validity of animal tests can, however, elicit a totally inappropriate response from scientists: "We must design and carry out better animal tests"! For example, attempts have been made to overcome the problem of species differences in drug metabolism by breeding and using animals who have been genetically altered to be more like humans. Using larger numbers of animals, and taking more samples, is another general approach that has been suggested for improving the predictive ability of the tests. All of these approaches are counterproductive and increase the burden on animals even more – this is not an acceptable way to solve the problem.

Animal toxicity tests should be replaced on animal welfare and ethical grounds, regardless of whether or not they are scientifically valid

When challenging the scientific validity of animal tests, it is essential to emphasise that ultimately the deficiencies of using animals as 'models' of humans can only be overcome by adopting fundamentally new approaches. The development and validation of advanced tests, based for example on *in vitro* methods,

[2] More information on the regulations governing the testing of different types of product, and the process of safety assessment, will be available in the RSPCA information paper 'Product safety regulations and testing on animals' (*available Autumn 2005*)
[3] APC (2003) *Review of Cost-benefit Assessment in the use of Animals in Research*. London: Home Office
[4] Nuffield Council on Bioethics (2005) *The Ethics of Research Involving Animals*. London: Nuffield Council on Bioethics

computer modelling, and ethically acceptable studies on humans, must go hand-in-hand with phasing out the animal tests.

1.1 Limitations of animal tests

There are many different animal toxicity tests, measuring different effects (from skin irritancy to birth defects and cancer), using different species, and giving different types of information. It is improbable that all these tests are totally invalid for all purposes, but there are many reasons why animal tests cannot be expected to predict the adverse effects of human exposure to a chemical with complete accuracy. Some of these relate to species differences in the way in which chemicals are absorbed, chemically altered, excreted, or bound to specific molecules within the body. Others are related to the specialised conditions used in the laboratory, lack of diversity in the animals used (genetic background, age, diet and so on), and the use of small sample sizes (small numbers of animals). In fact, similar considerations of diversity and sample size also cause serious problems in correctly interpreting clinical trials of new medicines on human volunteers.

For some examples of studies into the accuracy of animal toxicity tests, see Annex 1

The limitations of animal tests are widely acknowledged by scientists and regulatory authorities, and a number of studies have been done in an attempt to find out exactly how accurate or inaccurate various animal toxicity tests are. Some examples are given in Annex 1. However, a major problem with estimating the accuracy of animal tests is that the 'true' effect of each chemical (the actual effect of the substance on people) is often unknown, particularly if the tests suggest a severe hazard and the substance is only used under strictly controlled conditions.

In their defence, toxicologists say that the results of animal tests are only part of the process of assessing the risk of exposing people to substances, and that they need to be interpreted in the light of experience and by using existing knowledge of factors such as species differences in chemical metabolism. They claim that, by and large, toxicology has a good record of protecting human health, and that animal tests play an important role.

In fact, animal tests have become accepted as part of the standard way of predicting human toxicity without any formal attempt to show that the results they produce are valid.

1.2 Animal tests and adverse drug reactions

One of the areas where the success or failure of animal tests should be most obvious is in the testing of new medicines. Patients are deliberately exposed to drugs, and the actual effects of the drug should be easy to measure. These effects can then be compared to the predictions made on the basis of animal tests. Indeed, the high incidence of 'adverse drug reactions' (ADRs) reported by doctors has been taken as evidence that animal testing is alarmingly unreliable. However, this conclusion is unhelpful and all too easy to dismiss for several reasons. For example, drugs undergo extensive clinical trials in people before they are released onto the market. If the incidence of ADRs is high, then the validity of the clinical trials is at least as suspect as that of the earlier animal tests.

Part of the problem is the use of a relatively small homogenous sample of test subjects (whether human or animal), and failing to detect problems that might affect only a particular sub-population (for example groups with a specific genetic background). In fact, human trials generally involve larger numbers of subjects than the animal tests[5]. Another factor that must be taken into account is that a large proportion of ADRs are related to overdose, inappropriate use, unexpected interactions between drugs,

[5] Typically, up to 5,000 patients are treated in clinical trials. An equivalent repeated dose study in animals (e.g. a 1 month study) might use 160 rats and/or 32 dogs. In total, about 1,500 animals are used for all tests.

and the risky use of drugs in desperate clinical situations. Animal tests are not designed to prevent adverse reactions under these circumstances.

The largest study of the validity of animal tests on medicines has been carried out by the International Life Sciences Institute[6] (see also Annex 1), which looked at a large number of cases where drugs caused serious side effects during clinical trials. Since the purpose of animal tests is primarily to protect the health of the volunteers in these trials, the cases examined were all examples of where animal testing seemed to have failed. By re-examining the results of the animal tests, the authors showed that in 73% of cases they could find evidence of the side effects that later occurred in humans. Remarkably, they did not attempt to explain why the drugs proceeded to clinical trial regardless of this evidence. In fact, the study gets us very little further in deciding how accurate the animal tests really were. There is no indication from this analysis of how often the results of animal tests correctly predict that a drug is safe, nor how often drugs are dropped because of their effects in animals (in which case we will never know whether the drug would have caused toxicity in humans).

In conclusion, there is a pressing need for better, more comprehensive studies of the validity of animal tests that take proper account of the results of animal tests and how those results are interpreted and used. Until then, it will be very difficult to make real progress in the debate on validity.

Assumption 2: Are the results of animal tests used effectively?

This is a key question, since there is absolutely no justification for subjecting animals to pain, suffering, distress and lasting harm if the information obtained is not used (a) in the overall assessment of the risk a substance poses to human health, and (b) to underpin effective measures to manage that risk. No test should be done if it is superfluous; so appropriate decisions about the management of the risk to humans should result from each of the tests carried out.

In many cases, the safety testing of chemicals religiously follows regulations or guidelines that prescribe a set of tests, which is often called 'check-box' testing. This approach has been widely criticised because not all tests will give necessary or valuable information for all chemicals. The completion of a full set of tests, before the implications of the results of each are carefully considered, is an approach favoured by the authorities because it reduces administration. Industry, on the other hand, generally prefers an 'intelligent' approach, which tests first for those effects that are most likely to lead to stringent safety precautions or restrictions on use. **The RSPCA strongly supports this structured approach, not least because it allows the more imaginative use, and more rapid introduction, of non-animal methods.**

If the results of animal tests are not going to be used, then **the tests should not be done in the first place**

The need to develop and define an intelligent approach to safety testing is currently very acute because of recent changes to the European Union (EU) legislation on chemicals. A regulation on the Registration, Evaluation, Authorisation and restriction of Chemicals (REACH) came into force in 2007. Based largely on the traditional approach of prescribed sets of tests for chemicals, this Regulation could result in the testing of about 30,000 substances over the next ten years. Millions of animals could be used unless new approaches to safety assessment are rapidly developed and brought into use. **The RSPCA, and Eurogroup for Animal Welfare, have lobbied hard for a more thoughtful approach to gathering information under REACH, based on extensive sharing of information between companies, the use of alternative methods, and case-by-case consideration of the need for information for adequate risk management, given the intended uses of each chemical and the likely human exposure. We also support the tiered, stepwise testing schemes suggested by FRAME[7] and ECVAM[8].**

Even where animal tests are thought to be essential for safety assessment, their actual influence in ensuring the safe use of chemicals is often in doubt. The commercial, technological and economic importance of a

[6] Olsen, H., Betton, G., Robinson, D., Thomas, K., Monro, A., Kolaja, G., Lilly, P., Sanders, J., Sipes, G., Bracken, W., Dorato, M., Van Deun, K., Smith, P., Berger, B. & Heller, A. (2000). Concordance of the toxicity of pharmaceuticals in humans and in animals. *Regulatory Toxicology and Pharmacology* **32**, 56–67.
[7] Fund for the Replacement of Animals in Medical Experiments: Paper published in ATLA (Alternatives To Laboratory Animals) Vol 31, 7-19, 2003.
[8] European Centre for the Validation of Alternative Methods: Published as Supplement 1 of ATLA Vol 30, 2002.

chemical plays an important part in determining how rapidly and thoroughly the information provided by toxicity tests is acted upon. When industry is adamantly opposed to restrictions on a valuable chemical, they will strongly challenge the validity and accuracy of animal tests. The weaknesses of animal tests, as discussed above, are heavily stressed and exposed. This can lead to the seemingly endless repetition of animal studies, with modifications, that is totally inconsistent with the faith in the reliability of animal tests that is often expressed by regulatory authorities and toxicologists. **In an attempt to clarify the significance of a hazard, or the degree of risk posed by a chemical to humans, tests will be carried out in a widening range of animal species, with continual modification of the treatment and observation protocols (see Annex 2 for an example). This is a completely unacceptable situation.**

Industry is prepared to vigorously challenge the validity of animal tests - if the results indicate that the use of a commercially important chemical should be restricted

In the RSPCA's opinion, toxicity tests should not be conducted on animals unless it is clear beforehand exactly what action will be taken in the event of a particular outcome. Authorities should not demand animal tests if they do not have sufficient confidence in them to stand by the results and use them to enforce the necessary restrictions on the use of a chemical, without further testing.

Assumption 3: Are all chemicals worth testing?

Generally speaking, the authorities that regulate the marketing and use of chemicals do not assess the potential value of the substances they deal with. They are primarily concerned with ensuring their safe production and use. In the case of food additives, a technological need for a new additive must be demonstrated before it can be used, and for pharmaceuticals, demonstration of efficacy (evidence that a medicinal product is effective in the treatment of a disease) may be a necessary prerequisite for licensing. However, product licensing does not depend on the potential 'social' value, or need for a new product. For example, some pharmaceuticals are developed almost exclusively for commercial reasons, rather than to fulfil an urgent medical need. In particular, so-called 'Me-too' drugs are designed to do the same job as an existing drug but, being a different chemical substance, can have patent protection and provide greater profits for a particular company.

Authorities that control the use of animals for experimental purposes could – and should - include the likely societal value of chemicals and products as a factor when deciding whether to allow safety tests on animals. In the UK, the Home Office could consider the potential benefit of a chemical as a factor in the cost-benefit assessment required under the Animals (Scientific Procedures) Act 1986. In practice, this does not happen – the only benefit that is recognised (and it is usually regarded as sufficient) is that of protecting human and animal health and the environment against toxicity. **This means that animals are caused pain, suffering and distress to enable the marketing of products whose value is highly debatable. This is unacceptable to the RSPCA.**

Animals should not suffer to enable humans to have a never-ending array of products on the market

So, in general, the potential value of a product to society is not usually a factor that affects decisions to carry out safety testing in animals. The one exception to the rule is the ending of animal testing of cosmetics and cosmetic ingredients in the EU[9]. It has become widely accepted that the word 'cosmetic' implies 'trivial and unnecessary'. Cosmetics are seen by many people as 'vanity products', and animals should not suffer for human vanity. It should be clearly understood that the decision to end animal testing of cosmetics has not been taken on the assumption that their safety can be adequately assessed by other methods. The rationale is that we do not need new cosmetics, and that ceasing to develop them obviates the need for safety testing of new ingredients. There are sufficient good cosmetics ingredients that have already been passed as safe.

This principle could be extended to household products, such as 'new and improved' cleaners, because they only serve the trivial purpose of pandering to the desire for novelty, or for a misguided aspiration for total cleanliness. However, a major problem with ending animal testing on trivial or unnecessary products is the question of categorisation. The definition of a cosmetic product has caused difficulties in the past, since some

[9] The testing of cosmetics or their ingredients on animals was ended in the UK in 1997/8, and in the whole of the EU in 2009.

'cosmetics', such as sunscreens, have functions in protecting health. Defining 'household' products is even more difficult, and in both categories some products will have more real value than others. (Note, papers on the testing of cosmetics and household products on animals have been prepared by the Boyd Group, with substantial input from the RSPCA[10].)

Ideally, the potential value of each chemical should be assessed individually, but who should decide whether animal testing is justified or not? How should the benefits be defined and weighed? There are undoubtedly great difficulties in making judgements about the value of chemicals, which may have potential benefits of different kinds, for specific sectors of society, or particular individuals. Weighing these benefits against animal suffering would be contentious, and, on a practical level, would cut across many government departments.

The RSPCA was at the forefront of achieving the UK and EU bans on testing cosmetics on animals. A similar approach is taken for other types of inessential product, wherever they can be sufficiently categorised and defined. The Society advises the public of the animal testing carried out on new products and warns them of the potential consequences for animals of buying 'new and improved' products.

Assumption 4: Can toxicity testing be done without using animals?

It is now generally accepted (and it is a legal requirement throughout the EU) that animals must not be used in tests if there is an alternative method of obtaining the required information without using animals. Non-animal methods are, in fact, widely used in many stages of the assessment of the safety of various chemicals. However, their acceptance for regulatory purposes is limited and is dependent upon extensive standardisation and validation of each alternative method. This can take a very long time.

A concerted effort should be made to develop and implement humane alternative methods - **immediately**

The deficiencies of non-animal methods which are most often cited by toxicologists and regulators are (a) their inability to model the complex interactions that occur between tissues in an intact animal (or human), and (b) their inability to detect unexpected effects in any of the many tissues of the body. For example, the amount of a chemical that is needed to kill cells in a culture of human connective tissue could be very different from the dose which would be lethal to a rat or human, because the chemical may not be absorbed into the body, or because it acts specifically on liver or nervous tissue.

There are a number of ways in which these deficiencies can be overcome or negated. Computer methods and *in vitro* tests can be used to detect basic interactions between chemicals and biological target molecules. For example, chemicals can be tested for reaction with DNA *in vitro*, a method which has been used for many years to indicate whether a chemical is likely to cause mutations. This test tells us nothing about the dose that would cause birth defects or cancers in animals or humans, but it is used as an early indicator or 'screen' for substances that may be dangerous. Also, alternative tests can be combined into 'batteries' covering various aspects of a chemical's action. For example, effects on several important cell types can be tested *in vitro* simultaneously, and the results combined with mathematical predictions of how the chemical would be distributed in the body, and information from other *in vitro* tests on how the chemical is metabolised.

[10] *The use of Animals for Testing Cosmetics - A Discussion Paper from the Boyd Group*, July 1998, and *The use of Animals in Testing Household Products - A Discussion Paper and Statement of Principle*, December 2002: available at: http://www.boyd-group.demon.co.uk/

It is true to say that at present no alternative method, or combination of methods, can give a reliable indication of *all* the toxic effects of a chemical, and the doses at which the effects are likely to occur – but this information may not be needed. More research is needed to develop better non-animal tests for toxicity. **The RSPCA has supported such research and lobbies for more funding from the Government, the EU and other funding bodies. At the same time, the Society promotes the use of the available alternatives to the fullest possible extent, for example in its activities on REACH.**

In the long term, there are encouraging signs of progress towards replacing animals in toxicology. New methods, involving analysis of the way in which the activity the genes within individual cells are altered by toxic chemicals ('genomics') hold out promise for a radical overhaul of toxicology. The US National Toxicology Program has published a 'Roadmap for the future'[11] which envisages the "*development of, and gradual transition to, vastly improved and higher-throughput methods for predicting the toxicological impacts of environmental agents*". (Clearly they believe that vast improvement is necessary and possible.) This, they say, would involve a "*transition in methods from predominantly mammalian screens toward more* in vitro *systems*" and they also state that "*a central theme in all these activities is to be mindful of the animal resources needed and to strive to address the Three Rs effectively*".

Summary comment

The RSPCA believes that the four basic assumptions that underpin the current reliance on animal safety tests should be more vigorously challenged, because in many cases they simply do not hold up under scrutiny. **Hundreds of thousands of animals suffer and die in toxicology procedures every year, all because far too little priority is given to replacing animal tests within the regulatory system.**

Substantially more commitment to and financial support for the development and use of alternative approaches is needed from both government and industry. In addition, the overwhelming, monumentally bureaucratic regulatory system and the cautious and conservative approach of regulatory authorities responsible for assessing the safety of substances and products need to be challenged.

This situation is unacceptable on animal welfare, scientific and therefore ethical grounds.

Research Animals Department
Science Group
RSPCA
Aug 2010 (*original:June 2005***)**

[11] *A National Toxicology Program for the 21st Century: A Roadmap for the Future*. November 2004

Annex 1: Some examples of studies of the accuracy of animal tests

→ The Multicenter Evaluation of *In Vitro* Cytotoxicity (MEIC) study[1] was coordinated by the Scandinavian Society for Cell Toxicology. Data were collected from poisons centres and used to calculate the lethal blood concentrations of 50 chemicals to people. These concentrations were compared with the lethal doses of the same chemicals tested in rat and mouse acute toxicity tests. The correlation was relatively poor (R^2 = 0.62 to 0.65), and cell culture tests for cytotoxicity were at least as good.

→ An extensive interlaboratory study documented very high variability in eye and skin irritation test results among 24 labs[2]. Similarly, a comparison of 281 cases of accidental human eye exposure to various household products with rabbit eye test data for these products found little or no predictive value for the rabbit test (R^2 = 0.48)[3]. Investigators with the US Food and Drug Administration found no clear relationship between rabbit and human eye responses and concluded that the rabbit eye test is "plagued" by lack of reproducibility[4]. A comparison of data from rabbit tests and four-hour human skin-patch tests for 65 substances found that 45% of classifications of chemical irritation potential based on animal tests were incorrect[5].

→ Many criticisms have been voiced regarding rodent carcinogenicity (cancer) bioassays. Their reproducibility appears to be poor[6], but species differences in response are the greatest cause for concern. One study showed that cancer tests in rats were only 70% predictive for carcinogenicity in mice[7], and another that 46% of substances tested by the US National Toxicology Program were carcinogenic in rats but not in mice, and vice versa[8].

Given that rats and mice are more biologically similar to one another than either is to humans, it is reasonable to assume that rodent-human concordance is far less than 70%. It is now quite clear that many chemicals cause cancer in rodents by mechanisms that do not operate in humans (rodent-specific carcinogens)[9]. The concordance between rodent studies and human epidemiology studies is poor; one literature review found that 19 out of 20 probable human non-carcinogens tested in rodent bioassays gave positive results (induced cancer)[10]. Rodent cancer bioassays therefore produce an unacceptably large number of false positive results and, of greater public health concern, Salsburg[11] reported that rodent bioassays were capable of identifying only 37 % of a group of known human carcinogens.

→ An examination of the responses of 12 animal species to 11 groups of known human teratogens[11] revealed that positive predictivity of human birth defects may be as low as 40 percent in rabbits, which also exhibit a false negative rate of 40 percent. An unacceptably high rate of false positives is also a concern: Of 1,223 definite, probable, and possible animal teratogens, fewer than 2.3 percent can be linked to human birth defects.

→ A study of 140 drugs that caused unexpected human toxicity during clinical trials, suggested that for 71% of these drugs there was evidence of the relevant toxic effect in the results of the animal tests[13]. This is often quoted as evidence that the preclinical animal tests were accurate in 71% of cases. However, the study also indicates that rodent assays were only 43% predictive of human toxicity, the added predictive capacity being provided by additional tests in non-rodent species. It is also a fact that, in all 140 cases, the preclinical testing did not prevent toxicity occurring in human subjects, i.e it failed.

The poorest predictivity of animal tests was found to be liver toxicity (55% concordance) and hypersensitivity/cutaneous reactions (35%). Haematological and gastrointestinal toxicities had the highest rate of prediction by animal tests (91 and 85% respectively). The adverse effects studied were detected in human clinical trials and were only included if they were considered "severe", i.e. resulted in termination of drug development. Effects found after drug approval and general use were not included.

This analysis gives no indication of the value of animal studies in predicting severe toxicity, where compounds are not permitted to be given to humans at all. A very large number of candidate drugs never get to clinical trials - rightly or wrongly, we do not know. In other words, the rate of **true and false positive** predictions, based on preclinical studies, is not estimated. Similarly, drugs that showed no significant toxicity in either animals or humans are also excluded i.e. **true negative** predictions.

Annex 2: An example of how industry can challenge animal test results if it suits them

The toxicological evaluation of di(2-ethylhexyl) phthalate (DEHP) gives an interesting picture of how attitudes towards the reliability of animal tests may change when regulatory authorities, the chemical industry, and toxicologists are confronted with assessing the safety of an important industrial chemical which they want to continue to use despite concerns about its safety.

DEHP is a plasticiser used in vast quantities as a component of plastics. In rodent tests, it was found to cause liver damage and cancer, and atrophy of the testes. In spite of the apparent seriousness of these effects, the importance of the chemical meant that simply withdrawing it from use was not considered an option. A vast number of follow-up studies were done in an attempt to define the risks associated with human exposure to DEHP. There is now an enormous amount of information on the effects of DEHP in numerous non-human species, but firm conclusions about its safe use by people are still elusive.

The way in which DEHP causes cancer in the liver of rodents has been the subject of countless experiments. It now appears that the mechanism operates in rodents but not in primates, including humans (International Agency for Research on Cancer (IARC) - Summaries & Evaluations VOL.: 77 (2000) p. 41).

An evaluation of developmental and reproductive toxicity by the US National Toxicology Program (NTP) Center for the Evaluation of Risks to Human Reproduction (CERHR), in 2000, identified **41** animal studies of DEHP and its metabolites for developmental toxicity alone. Rats, mice and rabbits were administered various doses for various periods, by various routes. For reproductive toxicity, **68** animal studies were identified, using more species, including primates. These studies showed that the toxic response is variable in animal species, both in terms of the dose required to cause a response and in the nature of the effect. For example, rats and mice are the most sensitive species with the primary effects seen in the liver and testes; these effects are also observed at higher doses in hamsters and guinea pigs. In contrast, the liver and testes of cynomolgus monkeys, marmosets, and dogs appear insensitive to repeated exposure to DEHP. The report highlighted further research needs, so the uncertainty goes on and on.

Most industrial chemicals are tested only once, and the results are generally accepted by regulators and industry, especially when they tend to show that the chemical is not dangerous. But when the results cause industry a serious problem, they are challenged. Then we see how misleading or inconclusive animal tests can be.

References to Annex 1

1) Ekwall B *et al*. Overview of the final MEIC results: II. The *in vitro/in vivo* evaluation, including the selection of a practical battery of cell tests for prediction of acute lethal blood concentrations in humans. Toxicology *In Vitro* 13, 665-673 (1999).
2) Weil CS & Scala RA. Study of intra- and inter-laboratory variability in the results of rabbit eye and skin irritation tests. Toxicology and Applied Pharmacology 19, 276-360 (1971).
3) Freeberg FE *et al*. Correlation of animal test methods with human experience for household products. Journal of Toxicology: Cutaneous and Ocular Toxicology 3, 53-64 (1984) and 5, 115-123 (1986).
4) Koch W. Journal of Toxicology: Cutaneous and Ocular Toxicology 8, 17-22 (1989).
5) Robinson MK *et al*. (2002). *Food Chem Toxicol* 40, 573-592.
6) Gottmann E *et al*. Data quality in predictive toxicology: reproducibility of rodent carcinogenicity experiments. Environmental Health Perspectives 109, 509-514 (2001).
7) Zeiger E. Carcinogenicity of mutagens: Predictive capability of the Salmonella mutagenesis assay for rodent carcinogenicity. Cancer Research 47, 1287-1296 (1987).
8) Di Carlo FJ & Fung VA. Summary of carcinogenicity data generated by the National Cancer Institute/National Toxicology Program. Drug Metabolism Review 15, 409-413 (1984).
9) Cohen SM *et al*. Evaluating the Human Relevance of Chemically Induced Animal Tumors. Toxicological Sciences 78, 181–186 (2004)
10) Ennever FK *et al*. The predictivity of animal bioassays and short-term genotoxicity tests for carcinogenicity and non-carcinogenicity in humans. Mutagenisis 2, 73-78 (1987).
11) Salsburg D. The lifetime feeding study in mice and rats--an examination of its validity as a bioassay for human carcinogens. Fundamental and Applied Toxicology 3, 63-67 (1983).
12) Schardein JL. Chemically Induced Birth Defects, 2nd Ed. Rev. New York: Marcel Dekker (1993).
13) Olsen H *et al*. Concordance of the toxicity of pharmaceuticals in humans and in animals. Regulatory Toxicology and Pharmacology 32, 56–67(2000).