THE WAY FORWARD

ACTION TO END ANIMAL TOXICITY TESTING

REPORT COMPILED FOR THE BUAV BY
DR GILL LANGLEY MA, PhD (CANTAB), MIBiol, CBiol

BUAV
The British Union for the Abolition of Vivisection

European Coalition to End Animal Experiments

INTRODUCING THE EUROPEAN COALITION AND THE BUAV

The European Coalition to End Animal Experiments, (hereafter 'the European Coalition'), is Europe's leading alliance of animal protection organisations who have come together to campaign for effective and long-lasting change for laboratory animals. Formed in 1990 by animal groups across Europe, the Coalition now represents members across member states of the European Union plus a range of international observer groups drawing together organisations with a range of legislative, scientific and political expertise. Its membership base is currently expanding to include those member groups in EU accession countries. Current Observer/Member groups of the European Coalition are as follows:

Members: ADDA (Spain); Animal Rights Sweden; Animalia (Finland); BUAV (UK); BVTVG (Germany); Danish Society for the Protection of Laboratory Animals; DeutcherTierschutzbund (Germany); GAIA (Belgium); LAV (Italy); One Voice (France); SSPA (Switzerland) and Vier Pfoten (Austria). The Chair group of the European Coalition is the British Union for the Abolition of Vivisection (UK).

Observers: Animal Alliance of Canada; Doris Day Animal League (US); Dierenbescherming (Holland); Dr Hadwen Trust for Humane Research (UK); EFAP (Greece); IAVS (Ireland); International Fund for Animal Welfare EU; JAVA (Japan) and Eurogroup for Animal Welfare.

Established in 1898, the British Union for the Abolition of Vivisection (BUAV) is Europe's leading single-issue organisation campaigning to end animal experiments. The BUAV has an established record stretching over 100 years; we combine legal and scientific expertise, research skills, media liaison, public campaigning, undercover investigations and political lobbying in order to work effectively for an end to animal experiments and their replacement with modern and humane alternatives. The BUAV chairs the European Coalition to End Animal Experiments.

British Union for the Abolition of Vivisection (BUAV)
16a Crane Grove, London N7 8NN, UK.
Tel: +44 (0)20 7700 4888
Fax: +44 (0)20 7700 0252
Email: info@buav.org
Web: www.buav.org

ACTION TO END ANIMAL TOXICITY TESTING

Animal based toxicity tests cause massive suffering and are of dubious scientific value; their credibility is based on established use rather than reliability or predictive value.[i]

Testing programmes (such as the one proposed in the European Commission's White Paper on a Future Chemicals Policy) could drastically increase the number of animals used in toxicity experiments, or – with sufficient political will – can be used as an opportunity to bring non-animal tests into use.

This report demonstrates that animal tests can be replaced with modern, humane alternatives. It is aimed at all those whose influence and commitment is needed to help achieve this goal.

> For each test area, the animal test and its non-animal equivalent is described. Scientific failings of the established method are detailed, alongside action needed to bring new tests into use.
>
> The animal test regime described here is used to assess many substances, including cosmetic ingredients and other chemicals.
>
> Alternative methods described include tests already validated as well as those under development or already in use, although awaiting final validation by ECVAM (European Centre for the Validation of Alternative Methods).
>
> The description of each animal test gives an outline of the test method. It should not be forgotten that each test can cause pain and distress, often prolonged.

Under the conventional test regime, we estimate that up to 2123 animals are used to test each HPV (High Production Volume) chemical. It has also been estimated that a minimum of 12.8 million animals would be required to test 30 000 substances; this figure rises to over 50 million when offspring produced during reproductive studies are included[ii]

Animal tests are often slower than their non-animal equivalents. For example, the animal test for carcenogenicity takes five years. The non-animal SHE test could be validated in two years, and produces results in a few weeks (see page 23).

In 1993, MEPs in the European Parliament voted to end the testing of cosmetics on animals by imposing a ban on the sale of animal tested products, to be implemented in 1998. Non-animal alternative tests were promised but due to lack of funding (and provision for the ban to be delayed if validated alternatives were unavailable) development was not prioritised. Today, almost a decade on, we are still waiting for the new non-animal tests described here.

Further delay is unacceptable. The elimination of animal toxicity experiments, and their replacement with non-animal alternatives, is achievable. A well-funded and co-ordinated strategy to bring these non-animal tests into use is needed now.

i For each substance, up to twelve test areas are investigated, as required by legislation. In the majority of cases, the animal tests are accepted for use but not validated to modern standards; they have not undergone the validation process that is required for new non-animal tests.

ii IEH (2001) Testing Requirements for Proposals under the EC White Paper 'Strategy for a Future Chemicals Policy' (Web Report W6), Leicester, UK, Institute for Environment and Health (at http://www.le.ac.uk/ieh/webpub.html posted July 2001.

Contents

APPLYING THE NON-ANIMAL TEST STRATEGY

The current EU chemical testing review provides a unique and exciting opportunity for European institutions to reassess the value and weaknesses of the present chemical testing approach. Like other chemical testing programmes, it is an opportunity to develop a comprehensive new strategy for modernisation and improvement.

Instead of relying on the much criticized and highly uncertain testing methods of the past, the Commission should launch a new, coordinated and time-tabled research, validation and regulatory initiative based on accurate, efficient, cost-effective and humane methods. This will not only save animals' lives, but will also benefit environmental protection, human safety and consumer confidence.

Clearly, in order to identify hazardous chemicals and control them as soon as possible, a radically different testing process is needed. The only approach that combines practicality with humanity uses rapid non-animal tests to characterize a large number of chemicals in the minimum amount of time.

In vitro and computational methods can be combined in stepwise or decision-tree strategies tailored to each type of toxicity. Stepwise testing has already been accepted by, and is used within, the OECD (Organisation for Economic Co-operation and Development), the USA and the European Union. In a stepwise strategy using non-animal methods, testing progresses from quick and simple screening methods through tests specific to toxic mechanisms, to more sophisticated *in vitro* assays, where needed, which study target tissue effects. *In vitro* metabolism studies, combined with computer modelling of chemical absorption, distribution and excretion, permit extrapolation of *in vitro* results to the whole body situation.

Decision points about a chemical can be made at any stage of the testing strategy, so that not all the steps may be necessary for every chemical. For most types of toxicity the early steps are common, so that for each chemical the strategies can run in parallel at least up to a certain point.

Most *in vitro* tests provide results within days, rather than the weeks, months or years that are more typical of animal tests. Many of the tests proposed here are already familiar to industry, being used routinely as in-house screening methods and to understand mechanisms of toxicity. Some tests provide not only qualitative measures (is it toxic or not?) but also quantitative estimates of hazard (how toxic is it?).

Some of the non-animal tests proposed here have been validated and accepted by regulatory authorities. Others await final validation. Where appropriate, Priority Action is identified at the close of each section to highlight what 'next steps' the European Commission should take in order to implement a non-animal testing strategy.

Data on existing chemicals is needed urgently and because animal tests are too slow, often unreliable, expensive and resource-intensive, non-animal tests should be used exclusively. This will permit the effective and fast-track classification and control of likely toxic culprits. Chemicals that are borderline or suspect should be controlled on the precautionary principle pending tests with a fuller array of validated non-animal methods, where necessary, in the near future.

An essential prerequisite before any tests begin is mandatory data sharing as part of a worldwide search for existing data. This must include not just standard toxicology texts and databases, but data held by contract laboratories, chemical companies and by expert toxicologists. Records from hospitals, workplaces and poison information centres must be searched for details of reported effects of chemicals on humans. Data from animal tests conducted in the past to lower standards than today, can be integrated into the assessment and should not be a reason for conducting repeat animal tests.

THE WAY FORWARD

PRIORITY ACTION SUMMARY

Currently there exists a huge backlog of existing chemicals in use that have been suggested as requiring further testing. The appropriate control and regulation of these chemicals is an urgent priority. This co-ordinated, non-animal strategy should be utilised by the European Commission by implementing two approaches concurrently. Candidate chemicals can immediately undergo initial non-animal screens and early stage validated non-animal tests. *At the same time* the European Commission should immediately prioritise specific areas of validation, final validation or regulatory acceptance where required, for the potential non-animal tests specified in the strategy. By prioritising these non-animal tests for fast-track validation, the complete non-animal strategy would be complete for full implementation for the testing of existing chemicals (and new chemicals in the future) in the shortest possible time, and produce reliable results far more quickly and cheaply than by using the traditional animal-based experiments.

In summary therefore, in order to take advantage of the scientific, economic, practical and ethical benefits of this non-animal strategy for the testing of existing and new chemicals, the European Commission must at the same time implement some immediate priority action as follows:

1. immediate regulatory implementation of all validated alternatives
2. immediate prioritisation of potential alternatives through fast-track validation and regulatory acceptance
3. priority funding for alternatives in targeted areas, with clearly time-tabled processes for development, validation and acceptance

ANIMAL TEST | **EYE IRRITANCY**

ANIMAL:	**ADULT ALBINO RABBIT**
NUMBER:	**At least three per chemical tested.**
THE AIM:	**To assess the acute irritancy of a substance when applied directly to the eye.**
THE TEST:	**Single dose effects are monitored for up to 21 days. The other, non-tested eye serves as a control. In most cases no anaesthetic is given. An illustrated standard guide is used to score irritancy.**
SYMPTOMS:	**Cloudiness, reddening, lid swelling, ulceration, weeping eyes**

Criticisms

- marked changes are to be found in the structure of the cornea of the eye of the rabbit and that of the human. Rabbit corneal mean thickness being 0.37 mm as against 0.51 mm in the human. Also the rabbit cornea comprises 25% of the surface area of the eye in comparison to 7% in the human
- the Bowman's membrane, an important structural element of the eye, is six times thicker in the human eye than that of the rabbit
- the rabbit has a lower tear flow than that found in the human and so is less able to 'cry out' the test substance; this results in the longer residency of any material in the rabbit eye
- the scoring of any damage which results from the administration of test substances is highly subjective, there being significant variation both within and across laboratories. Swanston of Porton Down said "no single animal species has been found to model exactly for the human eye either in anatomical terms or in response to irritation" (1)
- eyelid differences may have an influence on the removal of any test substance
- the immune, physiological and genetic status of the animal will profoundly influence any response to irritancy and inflammation
- there are confounding effects of the volumes of test substance administered – there being no clear relationship between the rabbit eye response evoked by exposure to a fixed volume (0.1 ml) and that occurring in accidental human exposure (2)

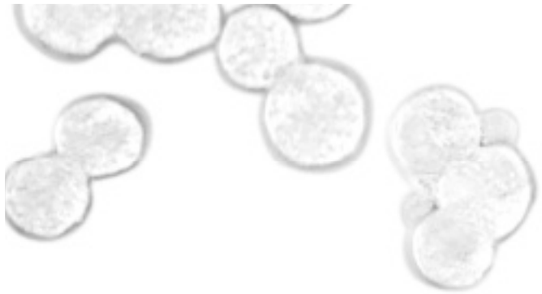
Alternative Strategy

After the collection and analysis of existing human or animal data, the following strategy can be followed (3):

STEPWISE TESTING STRATEGY FOR EYE IRRITANCY

STEP 1: Certain molecular structures predict irritancy, so structure/activity predictions can be made by computer systems to identify potential irritants.

STEP 2: Strong acids or alkalis are likely to be corrosive to the eyes and can be labelled as such without further testing.



STEP 3: It can be presumed that a chemical that causes severe skin irritancy will also be an eye irritant, and it can be labelled and managed appropriately on this basis.

STEP 4: The non-animal tests below have been validated to the satisfaction of national regulators (respectively Britain, Germany, France and the Netherlands) for identifying severe eye irritants. The isolated rabbit eye test and the HET-CAM test (Hen's Egg Test – Chorio-Allantoic Membrane) are able to rank chemicals according to whether they are mild or severe eye irritants. Under the OECD agreement on mutual acceptance of data, other countries will accept results from these tests. They include:

- Bovine Corneal Opacity & Permeability (BCOP) test
- HET-CAM test
- Isolated rabbit eye test
- Isolated chicken eye test

As well as these validated methods, the German Federal Institute for Health Protection of Consumers and Veterinary Medicine (BgVV) has stated (4) that a simple *in vitro* test of protein precipitation could further discriminate between chemicals that are moderate or severe eye irritants. Substantial protein precipitation in a test such as the Irritection system (formerly Eytex) would indicate irreversible eye irritancy.

On the basis of these non-animal tests, chemicals which are irritating or irreversibly damaging to the eye can be identified, labelled and controlled. Chemicals which are negative throughout the testing strategy are highly unlikely to be irritants.

THE WAY FORWARD

PRIORITY ACTION

ECVAM should assess the value of including an *in vitro* protein precipitation test such as the Irritection system.
ECVAM should review data on the BCOP, HET-CAM, rabbit eye & chicken eye tests, and ascertain whether they can be accepted by the EU on the basis of existing data, or whether further validation should be urgently undertaken.

Suggested timescale: 2 years maximum.

ANIMAL TEST | **SKIN IRRITANCY AND CORROSION**	
ANIMAL:	**ADULT ALBINO RABBIT**
NUMBER:	**At least three per chemical tested.**
THE AIM:	**To assess the toxicity of a chemical applied to the skin.**
THE TEST:	**A single dose is applied to an area where the fur or hair is removed and a shaved but untreated region close by is used as a control. Exposure is usually four hours. The irritancy is scored by checking it against the control patch of skin.**
SYMPTOMS:	**Reddening, swelling, inflammation and ulceration of the skin.**

Criticisms

- the anatomy of the skin and its cellular makeup varies in different species, thus any response will also vary between the species
- the rabbit, commonly used for irritancy tests, is a notoriously poor predictor of human irritation potential (5)
- this test procedure is fraught with difficulties because of the variations associated with the immune, physiological and genetic status of the rabbits used
- there are age differences in response both within and across species

Alternative Strategy

Chemicals that are corrosive to the skin can now be classified by validated test tube methods accepted by the EU (as below); consequently animal tests for skin corrosion are no longer permitted under Directive 86/609/EEC (6).

For skin irritation testing, the OECD has published a stepwise testing strategy in which *in vitro* data can be sufficient for chemical classification (but which permits an animal test as a final step) (7). The OECD strategy has several stages in common with a scheme prepared by the European Centre for the Validation of Alternative Methods (ECVAM) Task Force on skin irritancy (8), as follows, which does not necessitate the use of animal testing:

STEPWISE TESTING STRATEGY FOR SKIN IRRITANCY AND CORROSION

STEP 1: Certain molecular structures predict irritancy, so structure/activity predictions can be made with the assistance of computer models.

STEP 2: Strong acids or alkalis are likely to be skin corrosives and can be labelled as such without further testing.

STEP 3: Validated and approved *in vitro* methods for skin corrosion testing include the transepithelial electrical resistance test (using skin fragments *in vitro*), the human reconstituted skin model, or Corrositex.
Corrosives are identified and controlled accordingly. Chemicals negative for corrosivity go to the next stage or, in countries where volunteer studies are permitted, directly to step 5.

STEP 4: *In vitro* tests for skin irritancy: results are recognised by the EU and the OECD, but the tests have not yet been fully validated. A recent ECVAM validation study identified human reconstituted skin models as very promising tests (9). *Skin irritants can be predicted by this strategy although, where volunteer studies are permissible, human data provides the gold standard, as in step 5.*

STEP 5: Having identified corrosive chemicals, remaining substances can be classified on the basis of a 4-hour human patch test in volunteers. There is a standardised protocol (10) for this test, and the approach is supported by the ECVAM Task Force on skin irritation as well as being the subject of a draft OECD guideline. Such tests in humans provide gold standard data.

The strategy permits the rapid and accurate classification of chemicals for skin corrosion and skin irritancy. At present the OECD and EU schemes envisage an animal test as a last resort solely because human volunteer tests are not acceptable in some member states. To avoid animal testing completely, we suggest that, whilst the first four stages of the scheme can be conducted in those member states, step 5 could be contracted to another member state where volunteer studies are allowed.

THE WAY FORWARD

PRIORITY ACTION

This strategy can be implemented immediately. If some member states cannot accept the human (volunteer) skin patch test, the European Commission could also prioritise the final validation and acceptance of *in vitro* human skin model tests for skin irritancy. It should also work to gain regulatory acceptance throughout the EU for the human skin patch test.

Suggested timescale: 1 year

ANIMAL TEST | SKIN ALLERGY

ANIMAL:	**ALBINO GUINEA PIGS**
NUMBER:	**Minimum 17, though possibly as many as 30 per chemical.**
THE AIM:	**To assess the ability of a test substance to cause an allergic reaction in the skin.**
THE TEST:	**The test substance is delivered to the skin by surface application or by injection; hair is removed from the test site – usually the shoulder. Multiple doses are applied in order to cause (any) local reaction.**
SYMPTOMS:	**Reddening of the skin, skin cracking or peeling, swelling, inflammation and ulceration.** *NOTE: A mouse-based test which has less severe effects is likely to be introduced shortly by the European Union. This would use 16 mice, but is obviously not a non-animal method.*

Criticisms

- there are a wide range of test procedures – 15 different protocols which have varying dosage and application frequency – thus making extrapolation and meaningful comparison of results to humans very difficult; the subjective scoring for damage also compromises any comparison
- the test material is applied to an area of shaved skin; as the animal's skin is shaved and abraded, thus causing an initial 'insult' to the skin, this necessarily differs from normal human skin exposure where the skin does not have an initial insult. If there is a local irritation at the site of the injection then a weak allergy effect is easily overlooked
- the use of often large doses of a potential allergen does not mimic the ways in which allergies are triggered in the human
- the commonly used guinea pig as a model species shows the weaknesses associated with using any inbred species or animals where the genetic history is unknown. The comparison of data from guinea pig studies with other outbred or inbred strains is difficult to interpret
- a major question arises over the evaluation of allergen potency in different species, there being no evidence that potency in the guinea pig or mouse predicts potency effects in the human (11)
- there are wide differences in the various parts played by the immune and skin cellular components between 'model' animals and the human. Furthermore the microstructure of the test animal's skin is vastly different to the human
- the threshold in responding to a given concentration of any test substance is closely related to the health, immune status, diet and other exposures of potential irritating factors in that animal. The animal's immune, physiological and genetic predisposition also influences all the sensitization parameters

Alternative Strategy

Before any testing commences, existing data should be sought; these may include human test results most likely to be found in specialist hospital departments and in workplace health records.

STEPWISE TESTING STRATEGY FOR SKIN ALLERGY (SENSITIZATION)

STEP 1: Certain properties of chemical molecules, such as their ability to bind with protein, help predict their skin allergy potential. Therefore chemicals should be screened by computerised quantitative structure/activity programs, or rule-based systems such as DEREK (Deductive Estimation of Risk from Existing Knowledge).
Some sensitizers can be classified on this basis alone. Chemicals without structural alerts proceed to step 2.

STEP 2: Chemicals cannot cause allergic skin reactions unless they can penetrate the outer skin barrier to the living cells beneath. An *in vitro* skin penetration study (a validated method) using fresh skin fragments reveals this. It also provides information about metabolism of the chemical in the skin. This identifies chemicals which themselves are not allergens but which are metabolised to allergenic forms.
If the chemical and/or its metabolite have molecular structures associated with skin sensitization and can penetrate the skin, they can be assumed to be likely skin sensitizers (12). They can then be labelled and controlled, or further evidence obtained as follows:

STEP 3: To cause skin allergy a chemical must react with proteins. If it does not react with the protein human serum albumin in the test tube, it is very unlikely to be an allergen.

STEP 4: There are several *in vitro* tests using target cells that provide additional information, if needed. These include the use of human skin samples to test for the activation and movement of important cells called Langerhans cells (13) as well as test-tube studies of human dendritic cells.

A chemical which is positive in these tests is almost certainly a skin sensitizer, and vice versa.

THE WAY FORWARD

PRIORITY ACTION

***In vitro* tests in step 4 (the use of Langerhans cells for skin sensitization, and test-tube studies of human dendritic cells) require final method development and validation, and the Commission should prioritise this work.**

Suggested timescale: this could be completed within 5 years.

ANIMAL TEST | **ACUTE TOXICITY**	
ANIMAL:	**USUALLY RATS**
NUMBER:	**15-30 per chemical depending on protocol and route of administration (oral, dermal or inhalation).**
THE AIM:	**To assess the toxic effects on the whole body of a single dose of a chemical.**
THE TEST:	**The animals are housed for a minimum of five days in order for them to acclimatise. 14 days of observation is the norm and with oral dosing animals are fasted prior to testing. All animals are autopsied at the end of the test period. Any sex-specific response is noted.**
SYMPTOMS:	**Blood pressure changes, weight loss, excessive salivation, internal organ damage, breathing disturbances, convulsions, bleeding from eyes, nose or anus, pilo-erection, tremors, diarrhoea, coma and even death.**

Criticisms

- criticisms relating to the acute toxicity tests will also apply to the repeat dose toxicity tests and vice versa
- immune, physiological, genetic, sex and other health indexes are factors influencing the validity of test outcome. Similarly exposure to other chemicals encountered by the animal and the inbred status of animals of unknown genetic status introduces variability in outcome
- there are a number of significant species differences in the role of detoxifying organs and sites of toxic accumulation. Species show differences in their P450 enzyme profiles (in the liver) and the kidney clears substances in different ways and at different rates in animals used for testing and the human (14)
- the routes of delivery are also important variables and lead to wide variation between test species and even different laboratories. The time the body takes to clear away any given substance varies both between species and even from one strain to another of the same species
- the rates of detoxification and metabolism show variation in different animals – the body's preferred routes of elimination and the way the chemical changes once inside the body (biotransformations) show huge variation (16, 17 & 18)
- bloodstream protein-binding varies across the species used and also between the strains of the same animal – both acute and repeat dosing are handled in different ways (19)
- wide variation from the human situation is also encountered in which secondary metabolites are formed and where these are stored and cleared from the body – this influences toxic potential

ANIMAL TEST | **REPEAT DOSE TOXICITY**

ANIMAL:	**USUALLY RATS; DOGS MAY ALSO BE USED AS A SECOND SPECIES**
NUMBER:	**40-80 rats and/or 32 dogs.**
THE AIM:	**To assess the toxic effects on the whole body of repeated sub-lethal doses of a chemical.**
THE TEST:	**The animal is repeatedly dosed with a chemical over a period of 28-90 days. The substance is usually administered orally by gavage (force-feeding) but may also be administered dermally or inhaled. The animals are then killed and their tissues examined pathologically and biochemically. Sometimes a 'satellite' group of ten animals is treated with the highest dose level for 28 days. This procedure supposedly establishes a baseline against which the effects of lower doses can be measured.**
SYMPTOMS:	**Blood pressure changes, excessive salivation, anaemia, aggression, muscle weakness, hair loss, internal organ damage, pilo-erection, vomiting (in dogs), tremors, diarrhoea, coma and occasionally death.**

Criticisms

- criticisms of repeat dose toxicity tests also apply to the acute toxicity test and vice versa
- the time course of the metabolism and elimination of any test substance will influence the ways that repeat doses elicit a response. For example, in some animals but not others, the chemical accumulates in the body over time causing a more toxic effect, which will complicate any extrapolation to the human. The binding of the test substance to various organs and cells within the body will mean that there will be different distribution and concentration of the toxic substance in the internal organs of different species which will affect the interpretation of both single and multiple doses and the necessary extrapolation to the human
- in long term exposure studies there are serious problems in end-point evaluation (15)
- bloodstream protein-binding varies across the species used and also between strains of the same species – both acute and repeat dosing are handled in different ways (19)

Alternative Strategy

ACUTE AND REPEAT DOSE TOXICITY

These tests explore the effects of the chemical once it enters and circulates around the body. Existing data from animals or humans (eg. from poison information centres, hospitals, or factory records) should be integrated and analysed before new tests are considered.

An ECVAM expert workshop has published a stepwise strategy for acute toxicity based primarily on non-animal tests (20). Computer modelling, such as physiologically based biokinetic modelling, of the likely absorption, distribution and excretion of the chemical in the body is incorporated to enable an estimate of the likely safe dose in humans for each chemical, if any (21).

There is good evidence (22) that cell culture tests can be predictive of repeat-dose toxicity, as well as single-dose (acute) toxicity.

The first four steps of the ECVAM scheme are:

STEPWISE TESTING STRATEGY FOR ACUTE AND REPEAT-DOSE TOXICITY

STEP 1: Physical and chemical information about the chemical (such as dissociation constant, fat solubility, volatility).
This permits comparison with chemicals of known toxicity; and also enables computational techniques to predict uptake and distribution within the body.

STEP 2: Basic cell culture (cytotoxicity) tests identify chemicals that are toxic to most cells by non-specific processes. [See box 1 for details of MEIC programme (23)].
Highly toxic chemicals could be classified accordingly. Chemicals which do not appear to be generally toxic to cells would progress to the next stage.

STEP 3: *In vitro* data, including metabolism studies, are integrated with computer simulations of absorption, metabolism, distribution around the body and excretion. The outcome is a prediction of the likely chemical concentration and time course in target organs.
Toxic chemicals likely to persist in the bloodstream are classified at this stage. Target organs are also identified.

STEP 4: Using information from step 3, more specialised cell tests identify selective chemical toxicity to particular tissues, such as kidney, nervous tissue, vascular cells or heart.

Again, with the application of computer modelling, chemicals are classified on the basis of these tests. Any chemical that is negative throughout the strategy is unlikely to pose a toxic hazard. 'No-observed effect' and toxic dose levels can be derived from these studies.

THE WAY FORWARD

PRIORITY ACTION

The European Commission should prioritise the validation of the MEIC Programme or similar methods, especially for repeat dose testing. Step 4 tests for target tissue toxicity should be validated in consultation with ECVAM.

Suggested timescale: 5 years maximum

BOX 1 - MEIC PROGRAMME

A seven-year programme called the Multicentre Evaluation of In vitro Cytotoxicity (the MEIC programme) was organised by Swedish scientists to find out if *in vitro* tests could predict the toxicity of chemicals to humans. Dozens of laboratories around the world used the same set of reference chemicals in their *in vitro* tests, and the results have been published in a series of papers (23).

Out of 61 methods trialled, a combination of four human-cell culture tests has been identified as practical, cost-effective and highly predictive of human acute toxicity (as measured by lethal peak bloodstream concentrations obtained from cases of acute poisoning in humans). *The four in vitro tests predicted this human data better than rat and mouse LD50 (Lethal Dose 50%) tests predicted lethal doses in humans.*[1]

These *in vitro* tests are:

- Changes in protein content in a human cell line
- Changes in ATP in second human cell line
- Changes in cell shape and
- Changes in pH in a third human cell line

The organisers of the MEIC programme have called for a formal ECVAM validation study of the proposed four-test selection. If successful, the cell culture tests could provide quick and effective assessments of the acute toxicity of chemicals to humans, without recourse to animal tests.

An ECVAM expert workshop has recommended a tier-testing strategy for the classification of chemicals for acute toxicity. This combines cell culture tests of the kind identified by MEIC, with physico-chemical data; if necessary, a second tier would comprise metabolism of chemicals using liver cells in the test tube. A third tier would involve *in vitro* studies on specialised cells. The ECVAM experts stated that this entirely *in vitro* approach could classify toxic chemicals (20). The best tests for the second and third tiers should be selected by ECVAM.

On this basis thousands of existing chemicals could be screened quickly and cost-effectively for acute toxicity. Those with high predicted toxicity could be controlled or withdrawn on the precautionary principle pending fuller analysis when further *in vitro* tests, if necessary, are available.

1 Lethal blood concentrations and lethal doses are not identical endpoints. The *in vitro* tests and the rodent tests predict slightly different measures of acute toxicity but both are valid measures

ANIMAL TEST | **MUTAGENICITY**	
ANIMAL:	**RATS, MICE OR CHINESE HAMSTERS**
NUMBER:	**At least 40.**
THE AIM:	**To identify any mutagenic effects of the chemical either on the rapidly dividing cells of the bone marrow or on the nuclei of blood cells.**
THE TEST:	**The test material is administered either by mouth or by injection into the body cavity. Two control groups are used for each test substance and dose regime – one group receives no chemical and the second group receives a compound that is known to have effects on the genes. Single or multiple doses are used and tissue sampling is performed up to 48 hours after the dosing of the test material.**
SYMPTOMS:	**Animals undergo one or more injections into the abdominal cavity, which can be painful. Gene changes are the endpoint.**

ANIMAL TEST | **CARCINOGENICITY**	
ANIMAL:	**VERY YOUNG RATS OR MICE**
NUMBER:	**At least 400.**
THE AIM:	**To detect any cancerous changes as a result of exposure to a chemical.**
THE TEST:	**Dosing occurs as soon after weaning as possible. The usual route is orally, but substances can be delivered by skin painting and by inhalation. Routes are chosen to attempt to mimic the ways the human might be exposed to the chemical. The minimum toxic dose is assessed by monitoring the weight of the animal and the highest dose is that which shows a decrease in the bodyweight gain without inducing tumours. The outcome of exposure is assessed by blood sampling, pathological appearance and tissue and organ examination in order to detect cancerous changes. These tests can cost $2 million per chemical and take 5 years to complete.**
SYMPTOMS:	**Tumours, lethargy, nausea and even death.**

Criticisms

Mutagens are substances that cause genetic changes (mutations). Mutations are often one of the steps in the formation of cancer. For the sake of simplicity the following criticisms of tests for mutagenicity and carcinogenicity are combined. Although the animals' experiences are very different in the two types of tests, from a scientific point of view the criticisms are similar, since many chemicals which cause mutations will also cause cancers

- it is now recognised that cancer is not a single disease. Tumours in both humans and animals normally occur as a result of a number of events, only one of which may be a single chemical exposure – this is not taken account of in the animal tests used to assess mutagenicity or carcinogenicity
- many of the criticisms of toxicity testing are relevant to the tests for potential mutagens, especially the immune status and unknown genetic composition of test animals
- without detailed knowledge of the biochemistry of the test substance in the test species and the human, it is unclear whether there is penetration of the material to the bone marrow – thus false negative results are likely (24 & 25)
- because of possible difficulties of penetration of the test material to the bone

marrow, large doses (usually the maximum tolerated dose) are used. This makes extrapolation across species and to the human dubious in terms of usual exposure

- injection of test substances into the body cavity is not an appropriate route of administration for exposure to a range of chemicals. Combes has pointed out the various problems associated with this test of mutagenicity as applied to cosmetics (26)
- it is unlikely that the bone marrow is the primary target tissue for many potential mutagenic chemicals (26)
- metabolic rates vary with the size of the organism – thus rats and mice are high metabolic species compared to the human which is a relatively slow metabolic species (27)
- there are many natural 'scavenger' chemicals present in organisms – these act to clear any potentially harmful molecules which arise either within the animal or are ingested in some way. The rates at which depletion of one such 'scavenger', glutathione, in different animals occurs is a confounding factor in comparisons of likely cell damage produced by a genetically harmful substance that is being assessed (28)
- mutagenesis (genetic damage) and carcinogenesis are known to be both related to the effects of a highly active form of oxygen (29) and many animals used to assess carcinogenic and mutagenic potential have far less sophisticated DNA-repair mechanisms than that found in the human (30). It is not therefore surprising that small rodents are more prone to cancer than the human (31), and that making comparisons becomes a matter of guesswork (14)
- humans enjoy an unrestricted diet – thus their intake of natural and synthetic anti-oxidants is higher than animal 'models' fed controlled diets of low anti-oxidant composition – this makes it far more likely that laboratory animals are prone to various cancers (32)
- during their lives humans are exposed to potential carcinogens in a host of different doses and combinations. In test situations single substances are delivered and the toxic response will differ between single and multiple exposure to even the same chemical. (33) Therefore extrapolation to the complex situation encountered by people in their work or daily lives is highly problematic
- the physiological and immune status of the 'model' animals is significant to the outcome of any test procedure. Variation between animals in their diets, diabetic status, fat intake, trauma, stress and their housing will influence the outcome of test exposure (34)
- a suite of variations exist between the human and rodents in the major drug-metabolising enzymes of the liver, such as the cytochrome P450s. This influences the carcinogenic potential of various chemicals. Furthermore, the kinds of P450 that are found in the human depends upon diet, genetic constitution, smoking and alcohol consumption, and environmental exposure (14)
- species and strain variation within and across test rodent species indicates that many harmful chemicals are handled in the body in ways which do not follow a simple path. Mani et al (35) showed that benzene, which has potent toxicity and at low doses tumour-inducing potential, is dealt with by different inbred strains of commonly used mice in quite different ways depending upon the strain – the Fischer rat was different again. Differences appeared to depend upon the metabolic capacity of each strain and species. Similar conclusions are found when heterocyclic amines (a common carcinogenic organic molecule formed during various cooking procedures) are investigated. Turteltaub at al (36) found that these amines damaged genes in rodents, but more so in humans exposed to the same dose. Furthermore, there were substantial differences in the metabolites of the amine administration between the humans and the rat. The authors concluded that the data suggests that rodent 'models' do not accurately reflect the human response to heterocyclic amine exposure (36). Garner et al in their study of 1999 came to similar conclusions (37)
- humans tend to store many harmful soluble chemicals such as polychlorinated biphenyls (pcb) in fatty tissues – the toxicity being released upon metabolism; in the rat the body uses a quite different detoxification pathway, consisting of storage and breakdown machinery (the desaturase enzymes)
- variation is even found between different laboratories when tests of the same chemicals assumed to be cancer causing are undertaken in the same strain of rodent
- differences in the routes of administration will alter the response which any test substance evokes
- the tumour potential of any given chemical also shows species variation – in the rat saccharin administration has been shown to induce cancer of the bladder and this appears to involve a promoter protein acting in an alkaline pH urine – the protein is not found in humans and the urine of most people is acid (38)
- the finding of false negative and false positive results when comparison is made between test responses in the rodent and the human complicates comparisons. Here tests in rodents frequently misclassify chemicals as carcinogenic when they are not and vice versa
- animal tests of carcinogenic potential are both time-consuming and expensive – as has been shown this is not balanced by the reliability of the resultant data
- there are marked age differences within and across species in their responses to potential carcinogens and mutagens
- it has been found that dietary restriction in young mice will delay the appearance of tumours in response to a test with a suspected carcinogen (39 & 40). The relevance of this to humans is not clear
- there are wide species variations in the absorption of any chemical. Lethality varies between strains. For example, there is up to a five-fold difference in skin absorption rates between human and other species. The skin of humans is less permeable to many chemicals than that of rodents – such absorption data overestimates the risk of tumours (41). Major differences in absorption from the food can be traced to significant differences in the anatomy of the gastrointestinal system in different species (42)

Alternative Strategy

MUTAGENICITY

An estimated one-fifth of chemicals currently used in commerce are mutagenic (43). Chemically-caused mutations can lead to cancer or to birth defects.

Several *in vitro* methods for testing mutagenicity have been validated, authorised and used in regulatory toxicology for more than 20 years. These methods have long been in the OECD test guidelines and the corresponding EU annex. They are suitable for assessing single chemicals and chemical mixtures, and can identify mutagenic metabolites as well.

The US Environmental Protection Agency accepts data from two *in vitro* tests for genetic toxicity, without *in vivo* testing, in the US High Production Volume Chemical Challenge Program (44).

STEPWISE TESTING STRATEGY FOR MUTAGENICITY

STEP 1: Structure/activity relations and computer systems such as DEREK, COMPACT (Computer Optimised Molecular Parametric Analysis for Chemical Toxicity) and TOPKAT (Toxicity Prediction by Computer-Assisted Technology) can flag up likely mutagens.

STEP 2: Chemicals are tested in the three standard accepted *in vitro* assays: the Ames test, mammalian cell mutation tests and chromosome aberration assays, with and without metabolic activation.
Three positive results identify mutagenic chemicals and three negatives indicate lack of mutagenicity. In the event of mixed results, chemicals progress to step 3.

STEP 3: Chemicals with mixed results from step 2 can be additionally tested using tailored assay modifications, and by other methods (such as the *in vitro* micronucleus test which is currently undergoing validation in Japan, Europe and the US, or the COMET assay).
The actual likely risk posed by chemicals identified in these tests can be estimated using computational modelling techniques (see box 2).

This strategy permits the identification of chemicals that are and are not likely to cause mutations.

THE WAY FORWARD

PRIORITY ACTION

The European Commission should obtain definitive advice from ECVAM on the best selection of tests for step 3. The *in vitro* micronucleus test is currently undergoing validation; the European Commission should prioritise this and implement it as soon as it has been validated.

Suggested timescale: 1 year

Alternative Strategy

CARCINOGENICITY

Many carcinogens initiate cancer by causing mutations in the DNA, and these can first be identified by established *in vitro* tests (see mutagenicity strategy). However, having eliminated mutagenic chemicals, other chemical effects can also lead to tumour formation and a testing strategy for carcinogenicity must identify these.

Computer modelling of the likely absorption, distribution and excretion of the chemical in the body (see box 2) should be combined with *in vitro* studies of metabolism.

STEPWISE TESTING STRATEGY FOR CARCINOGENICITY

STEP 1: Basic properties of a chemical, including partition coefficient or dissociation constant, are relevant to carcinogenicity. Structural alerts or structure/activity relationships that link molecular structures to carcinogenic properties are also useful. These can be aided by computer modelling and expert knowledge systems (see Box 2).

STEP 2: Cell transformation tests such as the Syrian hamster embryo (SHE) cell transformation assay (See Box 3) have been endorsed by the authoritative International Agency for Research on Cancer, and by an ECVAM expert workshop.

STEP 3: Chemicals with mixed results for mutagenicity and step 2 tests (above) can be further submitted to mechanism-based *in vitro* study, such as the intercellular communication assay (48).

This strategy permits the identification of chemicals posing a carcinogenic risk, as well as chemicals that are not likely to cause tumours.

THE WAY FORWARD

PRIORITY ACTION

The European Commission should implement this strategy on the basis of the precautionary principle ie. a weight of evidence indicating carcinogenicity should lead to adequate control or withdrawal of the chemical. Final validation and acceptance of the SHE cell transformation assay in step 2 should be prioritised. On the advice of ECVAM, validation for the intercellular communication assay may also be prioritised.

Suggested timescale: 2 years

BOX 2 - COMPUTER MODELLING AND EXPERT SYSTEMS

There are a number of computer programs available which look at the structure of a chemical and its observed properties and extrapolate these to predict the behaviour of the chemical in the human body.

Available computer systems include:

- DEREK (Deductive Estimation of Risk from Existing Knowledge)
- COMPACT (Computer Optimised Molecular Parametric Analysis for Chemical Toxicity)
- TOPKAT (TOxicity Prediction by Computer-Assisted Technology)
- HazardExpert

TOPKAT was used recently by the US Environmental Protection Agency to prioritise water-disinfectant chemicals for their likely carcinogenicity (45).

COMPACT assesses the likelihood that a chemical will interact with liver enzymes and be activated by metabolism to a carcinogenic form (27).

HazardExpert is computer software which recognises carcinogenic structures in chemical molecules. A combination of COMPACT and HazardExpert systems correctly identified 85 per cent of 40 substances, carcinogens and non-carcinogens, including drugs and chemicals (27).

Physiologically based biokinetic modelling is particularly applicable in toxicity studies. PBBK models predict the likely absorption, metabolism, distribution round the body and excretion in the whole organism. Predictions can be made on the basis of standard published data on body compartments, combined with *in vitro* cell studies-derived information.

BOX 3 - SHE TRANSFORMATION TEST

When cells are 'transformed' they become cancerous. The test uses Syrian hamster embryo (SHE) cells in the test tube and takes several days to complete . Seventy-five chemicals that had been tested in rats and mice were put through the SHE test, to see if it could accurately distinguish carcinogens from non-carcinogens. The results were the same as rodent results in 83 per cent of cases, which is a very good performance (46).

An additional study with 19 chemicals, organised by the International Life Sciences Institute and due to be published in 2001, found an 89 per cent concordance with rodent results.

The SHE test was proposed to the OECD and a guideline was drafted in 1996; however, it has not progressed far through the OECD process. Yet the US Food and Drug Administration already accepts the value of cell transformation tests in drug development .

An ECVAM workshop published a report in 1999 on cell transformation tests as predictors of carcinogenicity (47) that stated that:

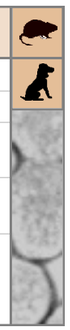
- Transformation tests were endorsed by the prestigious International Agency for Research on Cancer for their biological relevance to carcinogenesis, and as a logical approach for predicting the carcinogenic potential of chemicals ;
- The SHE transformation test can be highly predictive of carcinogenicity.
- A positive result in the Ames test and in the SHE test "should be considered to be strong evidence of rodent carcinogenicity of a chemical."
- It is justifiable to question why regulatory bodies have given a low priority to cell transformation tests.
- Immediate priority should be given to a validation study of the tests; the best method of performing them needs to be standardised.

Clearly, the SHE test is the key candidate for identifying cancer-causing chemicals. Final validation and standardisation of the test protocol is needed and could be achieved within a two-year period.

Combined with established *in vitro* mutagenicity tests, the SHE method would permit rapid identification of likely human carcinogens, with results in days or weeks instead of years. Thousands of animals' lives would also be saved.

ANIMAL TEST | **CHRONIC TOXICITY**

ANIMAL:	**RATS (DOGS MAY BE USED AS A SECOND SPECIES)**
NUMBER:	**160 rats (32 dogs).**
THE AIM:	**To assess the consequences of long-term administration of a chemical for significant periods of the animal's life span.**
THE TEST:	**The main routes of administration are oral and inhalation. The length of the study is at least 12 months and as much as 2 years.**
SYMPTOMS:	**Blood pressure changes, loss of appetite, aggression, restlessness, muscle weakness, excessive salivation, internal organ damage, pilo-erection, vomiting (dogs), tremors, bloody diarrhoea, coma and sometimes death.**



Criticisms

- see criticisms of acute and repeated dose toxicity tests (38). Again the physiological, immune and dietary status of the species chosen is paramount. Similarly the genetic predisposition of the 'model' animal is an area of ambiguity when interpreting the data from long-term studies of any kind (49). Also it is difficult to derive a reliable end-point for chronic toxicity
- in comparing the long term exposure of a potential toxic substance, how is scaling up achieved? In some cases it is known that organisms possess sensitive periods and so toxicity effects outside this period may be under estimated
- frequently differences in response stem from the sex of the test animal but the significance of this for humans is not always clear
- the test substance may damage detoxifying organs such as the liver and kidney – this may in turn render the test substance more toxic over time in those species with more susceptible livers or kidneys
- the natural history of exposure is crucial in deciding the long term toxic dose. Humans rarely receive one dose or the same size dose over a long period of time and so extrapolation from single or multiple doses in test animals is far too simplistic
- the route of administration is crucial in determining toxicity (see above). If test substances are given orally in animals having a vegetarian diet this route may well lead to the formation of harmful metabolites which are not found in the human (an example of such an artefact is where gut bleeding after penicillin administration in the guinea pig arises simply because the cellulose-digesting enzymes in the guinea pig mediate the bleeding – such an effect is absent in those non-herbivorous animals like the human) (50)
- the use of robust statistical tests is crucial and largely missing from many tests of the effects of chronic exposure
- there is little if any feedback from clinical and long-term human use which is fed into future test protocols
- any measurement of the effects of chronic or acute exposure to test substances must include subtle effects in order for meaningful extrapolation to the human. In tests of the effects of materials on the brain, no account is taken of changes in memory function, sensory effects and similar cognitive aspects. (51)
- the assessment of the impact of multiple routes is problematic. In the human scenario substances may be inhaled, eaten or delivered in fine particle form or to particular organs – this complexity is not built into simple tests with a 'model' animal. As has been indicated above, different species vary in how a test substance is detoxified, eliminated and how the health status of the animal influences the reliability of using the test data for human risk assessment
- chemical interactions between the test substance and its delivery medium are important. Chemicals which are dissolved in varying solvents will have very different effects from other closely related forms
- there are many uncertainties surrounding toxic effects which operate at different stages of growth in the human, where children may be far more susceptible to toxic chemicals than adults – use of an in-built 10-fold factor is assumed to cover safe intake thresholds but it is not clear that data assembled from animals can safely be used in this way (52)
- test substances may influence subsequent generations of offspring as was the case for use of stilboestrol (synthetic sex hormone). This chemical is now well known to be a potent cancer forming chemical but effects are not always found in the generation receiving the substance (49)

Alternative Strategy

In vitro methods for assessing chronic toxicity are not as well developed as for acute toxicity. Nevertheless, it is essential for practical as well as ethical reasons to test for long-term toxicity without using animals. Chronic toxicity testing in animals is a time-consuming (1-2 years, with intensive staff involvement) and costly exercise generating data which is difficult to interpret for humans.

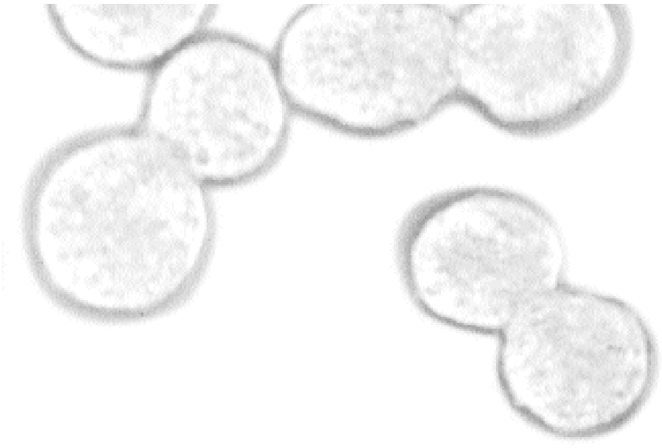
STEPWISE TESTING STRATEGY FOR CHRONIC TOXICITY

STEP 1: Structure/activity relationship screening using a computer system such as DEREK, is applied to identify likely toxic activity.

STEP 2: Basic cell toxicity tests conducted with longer chemical exposure are used to identify chemicals which cause general, non-specific toxicity to all cells. This information is also used in step 3.
Highly toxic chemicals can be classified at this stage.

STEP 3: *In vitro* data, including metabolism studies, are integrated with computer simulations of absorption, distribution and excretion of the chemical, as proposed by the ERGATT/CFN Integrated Toxicity Test Scheme (21). This yields predictions of the chemical concentration and time course in different body tissues, highlighting likely target tissues for each chemical. *Toxic chemicals that are likely to accumulate in the body can be classified at this stage.*

STEP 4: More specialised long-term methods (53) use cells from target tissues identified in step 3. These might be kidney, liver, nervous system or



vascular cells. “No-observed effect levels” of chemicals can be derived from these tests. For example, in a recent study (21) toxicity to nerve cell cultures of several chemicals, at a range of doses, was measured over 72 hours. This permitted a build-up of those chemicals, such as lindane, which accumulate in target tissues over time. Computational analysis of absorption, metabolism, distribution and excretion are applied here as in step 3.

Chemicals that cause toxicity to target tissues can be identified and threshold toxic doses quantified.

THE WAY FORWARD

PRIORITY ACTION

Further development of the step 3 *in vitro* metabolism and computer simulation studies and the step 4 target tissue tests should be prioritised by the Commission.

Suggested timescale: 5 years maximum

ANIMAL TEST | **TERATOGENICITY**	
ANIMAL:	**USUALLY RATS OR RABBITS**
NUMBER:	**At least 80 pregnant rats or 48 pregnant rabbits.**
THE AIM:	**To assess whether the test substance when ingested causes malformations in the embryo.**
THE TEST:	**A graduated dose or concentration is delivered to the pregnant female during the period of organ formation in the developing embryo. Three dose levels are given where the highest is sufficient to evoke minor changes in the mother (for example loss of weight). Tests are usually delivered by mouth and the embryos are killed and examined for gross or more subtle anatomical changes.**
SYMPTOMS:	**Female animals endure daily force-feeding by stomach tube throughout pregnancy. Females may experience poor weight gain, loss of appetite, nasal discharge, pilo-erection, hair loss, diarrhoea, dehydration and occasionally death. It is not established whether or not unborn animals who are damaged by a chemical feel pain.**

Criticisms

- the tests are designed to identify gross effects; more subtle clues about how the substance causes damage are therefore likely to be overlooked
- it is costly of animals and funds to run long-term exposure studies in order to evaluate the risk of various lengths of normal exposure in the human
- problems of relevance to the human situation arise when inbred strains of animals are used (54 & 55)
- there are differences in lifespan in various test species in comparison with the human – furthermore which comparison is the most appropriate for scaling i.e. is it body weight, length of exposure divided by the lifespan, or is it some mixture of the two?
- the estimation of suitable dosages is difficult to predict. In the case of thalidomide (56) different species are sensitive to different threshold doses
- many substances show teratogenicity in animals and not (as far as can be discovered) in the human. Aspirin causes malformations in the embryos of rats, mice, guinea pigs, dogs, cats and monkeys – even in high continued doses this does not appear to influence human embryos (56)
- there are vast differences in the structure and function of mammalian placentae. Substances pass through the placenta with different time courses and some cannot pass at all. Again there are differences in the carrier proteins and therefore the delivery of test material to the developing embryo
- the protecting effects of various natural or synthetic chemicals will also influence the ease with which teratogens can affect foetal and embryo development. Retinoic acid is dealt with by the body in varying ways and its toxicity can be either correspondingly small or large (57)

Alternative Strategy

Teratogenicity testing aims to identify chemicals that cause malformations or other damage to embryos in the womb. The standard test is labour-intensive and time-consuming.
Two non-animal tests (the embryo stem cell and the micromass test) have been undergoing final validation by ECVAM and successful results are imminent.

STEPWISE TESTING STRATEGY FOR TERATOGENICITY

STEP 1: Chemicals can be screened for molecular structures associated with teratogenicity by computer software such as DEREK and TOPKAT. TOPKAT was used recently by the US Environmental Protection Agency to screen water-disinfecting chemicals for their effects on the developing embryo (45). Chemicals with structural alerts can be prioritised for testing in the next stage.

STEP 2: The embryo stem cell test (EST), micromass test and post-implantation rat whole embryo test can all be used to identify teratogens (see box 4). *Computational modelling of absorption, metabolism, distribution to the embryo and excretion are applied, to generate information about dose levels likely to reach the embryo.*

By this strategy teratogens can be identified and controlled.

THE WAY FORWARD

PRIORITY ACTION

This strategy should be implemented immediately, as final validation studies of the embryo stem cell test and the micromass test have been completed. At the same time, any protocol standardisation, if needed, should be finalised and regulatory approval achieved. The Commission should further develop computer modelling techniques for step 2.

Suggested timescale: 2 years maximum

BOX 4 - TERATOGENICITY TESTING

Animal based teratogenicity testing uses a minimum of 80 pregnant rats or 48 pregnant rabbits, and is time-consuming and expensive.

During 1999-2001, three alternative tests were being validated by ECVAM. Two are based on rat embryos or embryo tissues maintained in culture, and one uses embryo stem cells:

- Embryo Stem Cell Test – EST
- Post-implantation rat whole embryo test
- Micromass test

The (EST) is the newest candidate alternative to animal tests. Stem cells survive in culture indefinitely and can develop into specialised cells such as heart cells, which can be seen beating in the test tube.

Test chemicals that cause teratogenicity prevent or limit the development of embryo stem cells into specialised heart cells in the culture dish (58). In the case of the two rodent embryo tests, in one (Post-implantation rat whole embryo test) whole embryos are studied in the test tube and in the other (Micromass test), cultures of limb cells from rat embryos are used.

So far, the ECVAM validation study (58) has revealed that each of the three tests can classify chemicals for teratogenicity or non-teratogenicity with an accuracy of 79-84 per cent, which is considered to be an excellent interim result.

All three alternative methods demonstrate very good potential for discriminating teratogenic and non-teratogenic chemicals, but only the EST is based on established cell lines with no requirement for a supply of animal embryos. This means the EST might be adaptable to automated high-throughput screening systems.

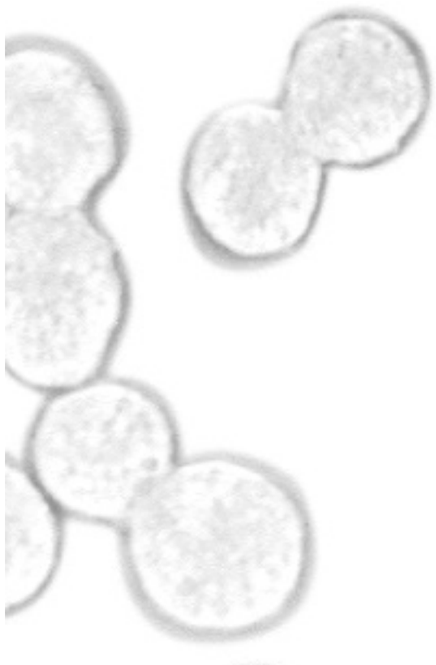
It seems highly likely that a valid alternative will emerge from the ECVAM study during 2001. Any protocol standardisation, if needed, as well as regulatory approval, could be achieved in two years.

ANIMAL TEST | **REPRODUCTIVE TOXICITY**

ANIMAL:	**RATS**
NUMBER:	**Approx 100 females (80 pregnant) and 40 males.**
THE AIM:	**To identify any effect of a chemical upon the male or female reproductive capacity.**
THE TEST:	**Graduated doses are administered (usually orally) to male and female animals during their reproductive cycles (sperm formation and absorption in the male and during two oestrous cycles in the female). Assessment is made of post-administration effects on fertility, pregnancy and maternal effects (feeding and nesting behaviour). Effects may be found in reproductive tissue, brain or secondary sexual systems.**
SYMPTOMS:	**Animals endure daily force-feeding by stomach tube. They may experience poor weight gain, loss of appetite, nasal discharge, pilo-erection, hair loss, diarrhoea, dehydration and occasionally death.**

Criticisms

- human and test species' reproductive systems and cycles are very different
- the influence of immune, physiological and dietary status on the interpretation of results of testing are fraught with the issues discussed in previous sections
- genetic constitution profoundly affects the reproductive toxicity of chemicals and this varies in humans and animals
- organs such as the testes and ovaries respond to the test substances differently in human and animal species
- a great many of the criticisms listed under chronic toxicity are also applicable here



Alternative Strategy

Several *in vitro* methods are used to study the mechanisms by which toxicity to reproduction occurs. Although some of the techniques are complex to perform, and have not yet been validated for regulatory use, nevertheless, a strategy is proposed which could identify chemicals likely to affect fertility, fertilization and implantation.

STEPWISE TESTING STRATEGY FOR REPRODUCTIVE TOXICITY

STEP 1: Physico-chemical properties of the substance are screened for structural alerts and information from basic cell tests may be relevant.

STEP 2: *In vitro* screening tests assess chemical effects on:
- Viability, motility, morphology and biochemistry of human sperm
- Sertoli cells in culture (cells important in sperm development)
- Ovarian follicle cells in culture
- Testosterone production by rodent Leydig cells *in vitro*
- Binding to cellular receptors such as for oestrogen and androgens.

Results from these tests can be extrapolated with computer modelling for an estimate of safe/toxic exposure levels. Chemicals with toxic effects can be labelled and controlled.

STEP 3: More complex tests can be done for chemicals that yield unclear results. For example, *in vitro* fertilisation can be studied; and embryo implantation can be modelled *in vitro* using a human cell line co-cultured with endometrial cells on a gel matrix (59).

THE WAY FORWARD

PRIORITY ACTION

***In vitro* screening tests in step 2 and the more complex *in vitro* tests in step 3 should be further developed and validated.**

Suggested timescale: 5 years

ANIMAL TEST | **TOXICOKINETICS**

ANIMAL:	**RODENTS AND SOMETIMES DOGS**
NUMBER:	**At least 8 of each species.**
THE AIM:	**To follow the time course of toxic effects and to answer several questions: Is the chemical quickly or easily absorbed from the gut or through the skin into the bloodstream? How long is it in the circulation? How is it excreted? Is it metabolised by the liver or other organs into a different chemical? Does it accumulate selectively in organs?**
THE TEST:	**Doses are either single or multiple. Healthy young animals are used and allowed to acclimatize to laboratory conditions for at least five days before tests are administered. Checks are made for sex dependency of response – the substance then given to the more responsive sex. Routes of administration are oral, inhalation or via the skin. After killing, the animal is examined for the accumulation of test substance in presumed target organs at different times from the start of the administration. Excretion and metabolism time courses are also followed.**
SYMPTOMS:	**Animals are isolated in metabolism cages, and some have tubes implanted into their bile ducts. Toxic effects include loss of appetite, lethargy, nasal discharge, pilo-erection, hair loss, diarrhoea, dehydration and vomiting.**

Criticisms

- see critique of toxicity tests for species and genetic problems associated with any interpretation. Similarly, the rates of detoxification and subsequent elimination are species- and strain-dependent and have serious problems of extrapolation to other species.
- bloodstream protein-binding varies across the species used and also between the strains of the same animal – both acute and repeat dosing are handled in different ways (19)
- wide variation from the human situation is also encountered in which secondary metabolites are formed and where these are stored and cleared from the body– this influences toxic potential
- species and strain variation within and across test rodent species indicates that many harmful chemicals are handled in the body in ways which do not follow a simple path. Mani et al (35) showed that benzene, which has potent toxicity and at low doses tumour-inducing potential, is dealt with by different inbred strains of commonly used mice in quite different ways depending upon the strain – the Fischer rat was different again. Differences appeared to depend upon the metabolic capacity of each strain and species. Similar conclusions are found when heterocyclic amines (a common carcinogenic organic molecule formed during various cooking procedures) are investigated, Turteltaub at al (36) found that these amines damaged genes in rodents, but more so in humans exposed to the same dose. Furthermore, there were substantial differences in the metabolites of the amine administration between the humans and the rat. The authors concluded that the data suggests that rodent 'models' do not accurately reflect the human response to heterocyclic amine exposure (36). Garner et al in their study of 1999 came to similar conclusions (37)

Alternative Strategy

The results of animal tests have to be converted into predictions for humans, but all these parameters can vary between species. The alternative approach is to combine *in vitro* studies, preferably with human cells, with computer simulations of the organ systems of the human body.

For example, the absorption of a chemical is studied using human gut cells in culture or, for skin absorption, isolated skin biopsy fragments. Likely metabolism by liver enzymes can be measured in the test tube using isolated liver cells or liver slices. *In vitro* methods can also show which proteins or cell receptors the chemical binds to. The physical and chemical characteristics of the chemical, such as its solubility in fat or fluid or its polarity, determine some of its toxic effects and can be measured in the test tube.

With these *in vitro* and molecular data, modern computer models (such as physiologically-based biokinetic models) can be used to provide answers to the toxicokinetic questions. The computer models provide estimates of how chemical effects vary at different doses, permitting effective chemical regulation. Not all these methods have been fully validated, but there are many years of experience with each of them and their usefulness has been accepted by regulators and companies alike. Drug companies routinely use these approaches in the development of new drugs.

THE WAY FORWARD

PRIORITY ACTION

The European Commission should implement this strategy using a precautionary approach ie. If the weight of evidence from the tests suggests health effects, a chemical should be controlled or removed. Validation of *in vitro* cell methods for gut uptake should be prioritised. Further development of computer models should also be prioritised by the Commission for wider application to chemicals testing.

Suggested timescale: 3 years

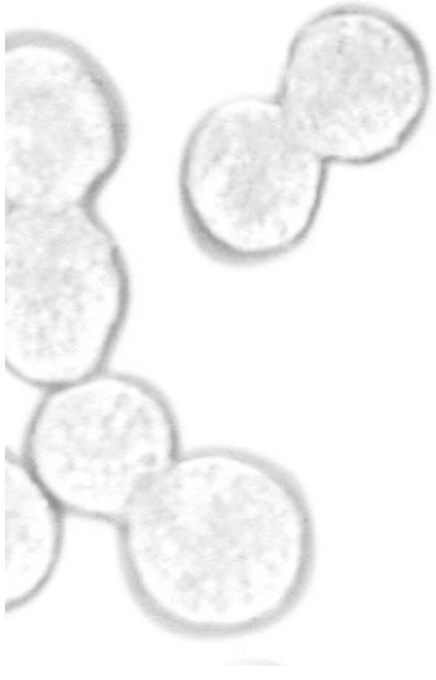
ANIMAL TEST | **ENDOCRINE DISRUPTORS**

There are no routine regulatory tests specifically for endocrine disruptors although new tests, both in animals and *in vitro*, are being developed; none has yet been fully validated.

Criticisms

A stepwise testing strategy is the best approach to testing for endocrine disruption, and in the standard model under development (eg. in the USA), first tier screening includes *in vitro* as well as animal tests. However, as none of the tests are validated it makes sense to focus on test-tube methods that are quicker and cheaper to perform.

- The Royal Society, in its 2000 report, admitted that animal tests cannot hope to explore the 'chemical cocktail' effect: "Using standard animal tests (acute toxicity tests) to evaluate these effects would be an extremely complex task with many potential problems" (60)
- genetic constitution profoundly affects the reproductive toxicity of chemicals and this varies in humans and animals
- organs such as the testes and ovaries respond to the test substances differently in human and animal species



Alternative Strategy

A chemical which has undergone testing for mutagenicity, carcinogenicity, reproductive toxicity or teratogenicity will already have plenty of data relevant to endocrine disruption. These existing data should be integrated so that as much information as possible can be gained without the need for further testing.

The proposed strategy could effectively be combined into the strategy for testing for reproductive toxicity described above.

STEPWISE TESTING STRATEGY FOR ENDOCRINE DISRUPTORS

STEP 1: Quantitative structure/activity relationships, assisted by computer analysis, can flag up chemicals likely to interact with estrogen and androgen receptors.
These results would help prioritise chemicals for further testing.

STEP 2: Androgen and estrogen receptor-binding using cell-free systems. The US Environmental Protection Agency is currently measuring the receptor binding of 500 chemicals to provide a database for endocrine disruptors. This approach could be used as a high-throughput system for very rapid screening of chemicals. Chemicals that increase or decrease the expression of key genes can also be tested *in vitro*.
Chemicals testing positive would be controlled accordingly, those testing negative could be further examined in step 3.

STEP 3: An *in vitro* test (the MCF-7 Focus Assay) uses human breast cancer cells grown in culture. In response to oestrogens, the cells alter their gene expression and the normal flat sheet of cells develops nodules of proliferation. Anti-estrogenic chemicals prevent this effect. The nodules (or foci) are easily counted to provide a measure of potency of the chemical and the test is currently undergoing validation.

THE WAY FORWARD

PRIORITY ACTION

In the absence of established tests of any kind, a non-animal strategy for potential endocrine disruptors should be prioritised and implemented on the precautionary principle.

The European Commission should work closely with the US Environmental Protection Agency and the OECD to co-ordinate the priority development and validation of non-animal methods. The *in vitro* MCF-7 Focus Assay is currently undergoing validation and should be prioritised by the European Commission.

Suggested timescale: Validated tests should be achieved within a maximum of 5 years.

References

1. Swanston D (1983) in: Animals and alternatives in toxicity testing. Balls et al (Eds). FRAME, London.
2. Koch, W (1989). J Tox - Cutaneous & Ocular Tox. 8:17-22
3. The first four steps of this strategy are the same as those recommended by the OECD in: OECD (1999). Detailed review document on classification systems for eye irritation/corrosion in OECD member countries. ENV/JM/MONO(99)4.
4. Gerner I, et al (2000). ATLA 28:11-28
5. Stitzel KA (2000). In: Progress in reduction, refinement and replacement of animal experimentation. M Balls, AM van Zeller and ME Halder (Eds). Elsevier 2000. pp 587-600
6. Directive on the Protection of Animals used for Experimental and other Scientific Purposes.
7. OECD (2000). Draft revised Guideline 404 (acute dermal irritation/corrosion) for testing of chemicals.
8. Botham PA et al (1998). ATLA 26:195-211
9. Fentem J et al (2001). Toxicology *in Vitro* 15:57-93
10. Griffiths HA et al (1997). Food and Chemical Toxicology 35:255-260
11. Basketter DA et al (1997). Food and Chemical Toxicology 35:417-425
12. Basketter DA et al (1995). Food and Chemical Toxicology 33: 1051-1059
13. Jacobs JJL et al (2001). Journal of Immunological Methods 247: 73-82; Verrier AC et al (1999) International Archives of Allergy and Immunology 120: 56-62
14. Parke DV and Ioannides C (1990). Progress in Clinical and Biological Research, 331: 215-230
15. Fox JG et al (1979) Food and Cosmetics Toxicology 17: 661
16. Dunne JF (1981) in: Textbook of Adverse Drug Reactions. D Davies (Ed)
17. George CF (1981). Clinical Science 60: 247-250
18. Garner RC et al (1999) Cancer Letters 143: 161-165
19. Marrs T (1988) in: Perspectives in basic and applied toxicology. Ballantyne B et al (Eds) John Wright and co, Bristol. p 288
20. Seibert H et al (1996) ATLA 24: 499 - 510
21. Blaauboer BJ et al (2000) Developments in Animal and Veterinary Sciences 31A: 525-536
22. Scheers EM et al (2001) Toxicology *in vitro* 15:153-161; Berlaimont V et al (1999) ATLA 27 (special issue): 415; Hartung T et al (1999) ATLA 27 (special issue): 415; McBride J (1999) ATLA 27 (special issue): 416
23. Clemedson C et al and Ekwall B et al (2000) ATLA 28 (Suppl 1); 159-234
24. Fielder RJ et al (1993) Human & Experimental Toxicology 12:189-198
25. Garner RC (1991) in: Animals and Alternatives in Toxicology: Present Status and Future Prospects. Ball M et al (Eds) FRAME, Nottingham
26. Combes RD (1994) in: conference proceedings Non-animal Alternatives in Cosmetics Testing – IFAW and Eurogroup for Animal Welfare, Germany.
27. Lewis DFV and Langley GR (1996). Mutation Research 369:157-174
28. Parke DV and Ioannides C (1990) in: Mouse Liver Carcinogenesis;Mechanisms and Species Comparison. Stevenson et al (Eds) Alan Liss, New York. pp47-63
29. Oberley LW and Oberley TD (1995) in: Drugs, Diet and Disease Volume 1. Ioannides C and Lewis DFV (Eds). Ellis Horwood, London. pp47-63
30. Hoeijmakers JHJ and Lehman AR (1994) in: Methods to Assess DNA Damage and Repair. Tardiff et al (Eds). John Wiley, Chichester. pp 57-82
31. Parke DV (1994) Environmental Health Perspectives 102:852-853
32. Smith M and Hepburn PA (1993) in Food, nutrition and Chemical Toxicity DV Parke et al (Eds) Smith-Gordon, London. pp119-127
33. Ross MH and Bras G (1965). Journal of Nutrition. 87:245-260
34. Ingelmann-Sundberg et al (1993) in Food, nutrition and Chemical Toxicity DV Parke et al (Eds) Smith-Gordon, London. pp147-160
35. Mani et al (1999) Toxicology and Applied Pharmacology. 159:83-90
36. Turteltaub et al (1999) Cancer Letters. 143: 149-155
37. Garner RC et al (1999) Cancer Letters. 143: 161-165
38. Monro A (1993) Regulatory Toxicology and Pharmacology 18: 115-153
39. Wolf NS (1995) letter in Science 270:1101-1102
40. Hart RW et al (Eds) (1995) Dietary Restriction: implications for the design and interpretation of toxicity and carcinogenicity studies, ILSI Press, Washington DC
41. Scott et al (1992) Fundamntal and Applied Toxicology 19:484-492
42. DeSasso JM and Jacobson CF (2001) Food and Chemical Toxicology 39: 209-228
43. Zeiger E and Margolin BH (2000). Regulatory Toxicology and Pharmacology 32:219-225
44. Green S (2000). Developments in Animal and Veterinary Sciences 31B: 1640
45. Moudgal CJ et al (2000). Toxicology 147:109-131
46. LeBoeuf et al (2000) Developments in Animal and Veterinary Sciences 31A: 107-111
47. Combes et al (1999). ATLA 27:745-767
48. Clothier et al (1995). ATLA 23:75-90
49. Festing MFW (1997) Nature 388: 321-322
50. Lawrence DR and Black JW (1978). The Medicine You Take. Croom Helm, London
51. Sullivan N et al (1993) Mutagenesis 8, 167-174
52. Renwick A et al (2000) Regulatory Toxicology and Pharmacology 31, 286-296
53. Berlaimont V et al (1999) ATLA 27 (Special Issue): 415;Hartung T et al (1999). ATLA 27 (Special Issue): 415; McBride J (1999) ATLA 27 (Special Issue): 416
54. Lewis P (1982). Archives of Toxicology (Suppl 5) 195-196
55. Neubert D (1989) in: Advances in Applied Toxicology AD Dayan and AJ Payne (Eds) Taylor & Francis, London
56. European Parliament Briefing, BUAV, Number 2, 1990
57. Nau H (1994) in EUROTOX Newsletter 17: 8-9
58. Genschow E et al (2000). *In Vitro* Molecular Toxicology 13:51-66
59. Genbacev O et al (1993). Reproductive Toxicity 7 (Suppl 1) 74-94
60. Royal Society (2000). Endocrine Disrupting Chemicals. p6. Royal Society, London