ALTERNATIVE (NON-ANIMAL) METHODS FOR CHEMICALS TESTING: CURRENT STATUS AND FUTURE PROSPECTS

A REPORT PREPARED BY ECVAM AND THE ECVAM WORKING GROUP ON CHEMICALS

Edited by Andrew Worth and Michael Balls

**European Centre for the Validation of Alternative Methods
Institute for Health & Consumer Protection
European Commission Joint Research Centre
21020 Ispra (VA), Italy**

THIS IS THE FINALISED DRAFT, WHICH HAS BEEN SUBMITTED TO *ATLA* FOR PUBLICATION

May 2002

LIST OF CONTRIBUTORS

Members of the ECVAM Staff

Michael Balls
Alessia Bogni
Susanne Bremer
Silvia Casati
Sandra Coecke
Chantra Eskes
Pilar Prieto
Enrico Sabbioni
Andrew Worth
Valérie Zuang

Members of the ECVAM Working Group on Chemicals

Martin Barratt	Marlin Consultancy, Bedford, UK
Bas Blaauboer	Institute for Risk Assessment Sciences, Utrecht University, Utrecht, The Netherlands
Philip Botham	Syngenta, Macclesfield, UK
Robert Combes	Fund for the Replacement of Animals in Medical Experiments (FRAME), Nottingham, UK
Johannes Doehmer	GenPharmTox BioTech AG, Munich, Germany
Julia Fentem	Safety and Environmental Assurance Centre, Unilever, Sharnbrook, UK
Manfred Liebsch	ZEBET, Berlin, Germany
Horst Spielmann	ZEBET, Berlin, Germany

Other External Partners

David Basketter	Safety and Environmental Assurance Centre, Unilever, Sharnbrook, UK
Richard Clothier	School of Biomedical Sciences, University of Nottingham, Nottingham,UK
Mark Cronin	School of Pharmacy & Chemistry, Liverpool John Moores University, Liverpool, UK
Per Garberg	Biovitrum, Stockholm, Sweden
Nicola Gilmour	Safety and Environmental Assurance Centre, Unilever, Sharnbrook, UK
Rosalind Hanway	Health & Safety Executive, Bootle, UK
Ian Kimber	Syngenta, Macclesfield, UK
Camilla Pease	Safety and Environmental Assurance Centre, Unilever, Sharnbrook, UK
Walter Pfaller	Institute of Physiology, Innsbruck University, Innsbruck, Austria
Barry Phillips	Royal Society for the Prevention of Cruelty to Animals, Horsham, UK
Vera Rogiers	Vrije Universiteit, Brussels, Belgium
Helmut Tritthart	University of Graz, Graz, Austria
Erik Walum	Biovitrum, Stockholm, Sweden

OVERVIEW

The principal aim of this report is to summarise the current status of alternative tests for contributing to the assessments of the potential toxicological (human health) effects that are currently required by European Union (EU) legislation on chemicals. These assessments are also likely to be required when the EU Chemicals Policy is implemented. The effects covered are acute lethal toxicity; dermal and ocular irritation and corrosion; skin and respiratory sensitisation; target organ and target system toxicity; genotoxicity and carcinogencity; reproductive toxicity; and the biokinetic endpoints of absorption, distribution and metabolism. The emphasis of the report is on methods that can be used to replace or partially replace existing animal tests. For each toxicological effect, the report identifies alternative methods that can be used immediately, either because they have been scientifically validated as definitive tests, or because they are considered to be sufficiently well-established for use in the prioritisation of further testing. For most of the toxicological effects considered, the incorporation of alternative methods into tiered testing strategies offers the most promise of replacing, reducing and refining the use of animals, without compromising the protection of human health. In some cases, such as dermal and ocular irritation and corrosion, tiered testing strategies can be implemented immediately.

The second main aim of the report is to make recommendations for the further development and validation of alternative tests and testing strategies, with emphasis on the expectations that could realistically be met in the short-term (by the end of 2003), medium-term (end of 2006), and long-term (end of 2010), if sufficient and appropriate human and financial resources were made available. This time-frame takes into account the duration of the forthcoming Sixth and Seventh EU Framework Programmes.

The report was compiled by the European Centre for the Validation of Alternative Methods (ECVAM), on the basis of contributions from a large number experts in the fields of toxicity testing and alternative methods, including members of the ECVAM staff, the ECVAM Working Group on Chemicals and associated Working Groups, and other experts in the field.

Chapter 1 summarises the background to the future Chemicals Policy, by referring to the main recommendations of the European Commission White Paper, to the conclusions of the European Environment Council, and to the opinions of the European Parliament and Economic and Social Committee. This chapter also explains the role that ECVAM has played in the formulation of the policy, and the role that ECVAM is expected to play in its future implementation.

This report refers to a number of completed and proposed prevalidation and validation studies. Therefore, Chapter 2 provides background information on the principles of validation, and on the way in which these principles are applied by ECVAM, which has the duty of coordinating the validation of alternative tests at the EU level. This chapter also explains the important concept of the prediction model, which enables the data generated by alternative tests to be interpreted in terms of potential toxic hazard either in whole animals or humans.

Chapter 3 describes the risk assessment process, in which the use of alternative tests is becoming increasingly significant. The concepts of hazard prediction and risk assessment are explained, and the possible uses of alternative tests in the different components of the risk assessment process are described. In addition, reference is made to the procedures of reverse risk assessment and read-across, which can play a useful role in reducing the need for animal testing.

The general aim of Chapters 4-10 is to address the current status of alternative tests in different areas of toxicity testing. Each chapter includes an assessment of the short-term, medium-term and long-term possibilites for the development and validation of alternative tests.

Chapter 4 deals with acute lethal toxicity testing, which is conducted as part of acute toxicity testing, and as a basis for making decisions on the need for further testing. There is a good medium-term prospect for the (partial) replacement of the classical LD50 test and other equivalent animal tests, since a validation study on the use of basal cytotoxicity tests (for certain kinds of chemicals) is to be initiated in 2002, under the auspices of the US Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) and ECVAM, taking into account long experience in the use of *in vitro* cytotoxicity tests and a number of evaluation studies. In this chapter, the outline of a tiered testing approach is also recommended.

Chapter 5 summarises the status of alternative tests for acute dermal and ocular irritation and corrosion. Significant progress has already been made with respect to skin corrosion, since guidelines based on two *in vitro* tests have been accepted at the EU level as replacements for the Draize skin corrosion test. Several alternative tests for skin irritation have been evaluated in an ECVAM prevalidation study, which is a high priority for follow-up in 2002. For eye irritation, it has not proven possible to formally establish the scientific validity of one or more replacement tests, applicable across the full range of eye irritation potency, despite a significant validation effort in this area. Thus, eye irritation remains a priority area for further research, test development and validation.

Chapter 6 covers dermal and respiratory sensisation, for which non-animal methods have yet to be validated. Considerable progress has been made in the development of the understanding of the biological mechanisms underlying the sensisation response, and a number of promising computer-based and *in vitro* systems are in the course of development. Thus, sensisation is also a priority area for further research, test development and validation. In the meantime, an animal-based (refinement and reduction) alternative, the murine local lymph node assay (LLNA), is regarded as scientifically valid for skin sensitisation testing.

Chapter 7 addresses the various biokinetic endpoints that relate to the absorption, distribution and metabolism of chemicals. Biokinetic parameters are important determinants of systemic toxicity, and their incorporation into biologically-based kinetic models will ultimately provide efficient tools for risk assessment purposes. This chapter refers to number of promising computer-based and *in vitro* models of membrane permeability that are considered ready for validation or regulatory acceptance and application. For the assessment of metabolism, both computer-based

and *in vitro* approaches are discussed, and a tiered approach is recommended, based on the sequential analysis of: a) enzyme pathways, including enzyme activation and inhibition; b) induction effects; and, where appropriate, c) enzyme polymorphisms.

The large and diverse area of target organ and target system toxicity is tackled in Chapter 8, which refers to nephrotoxicity and neurotoxicity by way of illustration. For effects such as these, it is clear that testing strategies, based on the use of complementary *in vitro* endpoints, will need to be designed and evaluated, and this will require a substantial and long-term effort. For neurotoxicty testing, a tiered approach based on the sequential assessment of basal cytotoxicity and neurospecific endpoints is recommended, and is approaching readiness for prevalidation. Chapter 8 also describes developments in the field of *in vitro* repeat-dose toxicity testing, which is recommended as a priority area for future research.

Genotoxicity and carcinogencity are covered in Chapter 9, which presents the outline of a tiered testing strategy based on computer-based models and *in vitro* tests for detecting point mutations, clastogenicity, aneuploidy and non-genotoxic carcinogenicity. It is argued that genotoxic chemicals should automatically be regarded as carcinogenic, to avoid the need to conduct time-consuming and expensive rodent studies. Efforts should be focused on the development and validation of *in vitro* methods for the detection of non-genotoxic carcinogens, and on the validation of computer-based models for predicting genotoxicity and carcinogenicity.

Chapter 10 discusses alternative methods for reproductive toxicity testing. This is another area in which testing strategies will need to be based on a number of critical and complementary *in vitro* endpoints, to provide a means of assessing the effects on different components of the reproductive cycle. Encouraging progress has been made in the area of developmental toxicity, since three alternative methods for predicting embryotoxic potential have been endorsed by the ECVAM Scientific Advisory Committee as scientifically valid and ready for consideration for regulatory acceptance and application. Future efforts should focus on the further development and validation of alternative tests for assessing other critical aspects of male and female reproductive cycles, and on the design and evaluation of testing strategies that incorporate these tests.

Chapter 11 discusses the current status of animal and non-aniimal methods for the detection of so-called "endocrine disruptors", i.e. chemicals that alter the function of the endocrine system and cause adverse health effects. International activities initiated by the US Environmental Protection Agency (EPA) and the Organisation for Economic Cooperation and Development (OECD) are described, with particular reference to three animal tests: the uterotrophic, Hershberger and an Enhanced repeat-dose protocol (OECD Test Guideline 407). In addition, various non-animal approaches are described, including computer-based and *in vitro* methods. This is an area in which more research and development ar required, so that suitably developed and validated non-animal tests can eventually be incorporated into tiered testing strategies that will reduce, refine and replace the use of animals.

Finally, Chapter 12 provides a summary of the current situation in relation to each toxicological endpoint, and brings together the main conclusions and recommendations presented elsewhere in the report. A distinction is made between

prospects for validation on the one hand, and recommendations for research and test development on the other. In addition, a number of recommendations are directed at specific organisations or expert groups, and some general recommendations are made for the use of alternative methods in the future EU Registration, Evaluation and Authorisation (REACH) system.

CONTENTS

1.   BACKGROUND

1.1   The White Paper on a Future Chemicals Policy

The legal requirements for the testing of chemicals in the European Union (EU) are laid down by four main pieces of legislation:

1) the "Dangerous Substances Directive" (*Council Directive 67/548/EEC*; Anon, 1967) and its subsequent amendments and adaptations to technical progress;
2) the "Dangerous Preparations Directive" (*Council Directive 88/379/EEC*; Anon, 1988) and its subsequent amendments and adaptations to technical progress;
3) Council Regulation (EEC) No. 793/93 on the evaluation and control of risks of existing substances (Anon, 1993); and
4) the "Limitations Directive" (*Council Directive 76/769/EEC*; Anon, 1976) and its subsequent amendments and adaptations to technical progress.

The current system of EU chemicals legislation distinguishes between "existing substances", i.e. chemicals that were declared to be on the market on or before 18 September 1981, and "new substances", i.e. chemicals that have been placed on the market since that date. A total of 100,195 existing chemicals are listed in the European Inventory of Existing Commercial Substances (EINECS), and approximately 3000 new substances have been notified since 1981.

It is now recognised that there are a number of weaknesses in the current system:

1) new and existing chemicals are not subject to the same testing requirements, which means that there is a lack of knowledge about the potential danger represented by many existing substances, which make about 99% of the total volume of chemicals on the market;
2) the current process of risk assessment is much too slow, with only a handful of existing chemicals being assessed each year; and
3) resources are concentrated too much on the assessment of new chemicals, which make up only about 1% of the total volume of substances on the market.

With a view to overcoming the weaknesses in the current system, on 13 February 2001, the European Commission adopted a *White Paper on a Future Chemicals Policy* (Anon, 2001a). The main objective of the new policy is to ensure a high level of protection for human health and the environment, while ensuring the efficient functioning of the internal market and protecting the competitiveness of the EU chemical industry.

The White Paper proposes to harmonise the testing requirements for new and existing substances, by introducing a new system for the Registration, Evaluation and Authorisation of new and existing chemical substances, known as the REACH system. Registration will require producers, importers and downstream users to deposit information with a central authority (probably the European Chemicals Bureau, a unit of the European Commission's Joint Research Centre), having conducted such additional testing as may have been necessary. The evaluation of these data will be managed by Competent Authorities in the Member States, and may

lead to further test requirements for substances produced in quantities of more than 1 tonne/enterprise/year. Authorisation will apply to chemicals of very high concern, to be identified during the evaluation process, irrespective of their production volume. It is proposed that authorisation will apply to carcinogenic, mutagenic and reprotoxic (CMR) substances (categories 1 and 2), and to so-called "persistent organic pollutants" (POPs). It is also proposed that a system of restrictions will apply to all other substances of concern, such as persistent bioaccumulative and toxic substances (PBTs), and very persistent and very bioaccumulative substances (VPVBs).

It is estimated that implementation of the new chemicals policy will result in the need for the further assessment of up to 30,100 existing chemicals, which are currently marketed in volumes greater than 1 tonne per year (t.p.a), and for which some essential human health and ecotoxicological data are lacking. The proposed schedule for testing depends on the production/importation volume of the chemical (Table 1.1).

The amount of testing required will be triggered partly by the production/importation volume (Table 1.2). For example, it is further proposed that chemicals with volumes in the range 1-10 t.p.a. should be tested with *in vitro* methods *alone*, which means that a set of appropriate alternative methods should be developed, validated and accepted by 2008. It should also be noted that chemicals produced/imported in amounts higher than 10 t.p.a. are not necessarily excluded from *in vitro* testing. In addition, the testing of chemicals of particular concern, such as CMRs and POPs, may be required, even if they are marketed in volumes less than 1 t.p.a.

1.2   Follow-up to the White Paper

The White Paper was discussed by the Environment Council on 7 June 2001 (Anon, 2001), when it was concluded, *inter alia*, that:

> "Animal testing should be limited to the level necessary to deliver the objectives of the strategy, including a high level of protection for human health and the environment. Industry should make all existing data available to avoid duplication of testing. Mechanisms are needed to ensure that unnecessary testing requirements are avoided. Adequate resources must be provided for research, development and validation of globally accepted test guidelines for alternative *in vitro* test methods, so that work can be accelerated at all levels. Activities under the new Framework Programme for Research should consider these requirements among its priorities. In addition to promoting this issue in ECVAM (European Centre for the Validation of Alternative Methods), the Community should play a more active role in the OECD, to encourage wider adoption of validated, alternative, non-animal testing methods." (Council Conclusion 23)

The White Paper has also been discussed in the European Parliament (EP) and in the Economic and Social Committee (ESC). On 15 November, a plenary session of the EP adopted a report on the White Paper produced by Mrs Schörling MEP. The Schörling report (Anon, 2001b) sets out a whole series of rules which would restrict the number of animal tests and foster the development of alternative methods. On 17 October 2001, the ESC adopted an opinion on the White Paper, in which it endorsed the commitment to promote non-animal testing (Anon, 2001c).

In order to prepare legislative proposals for the implementation of the White Paper, the Commission has established eight Working Groups, to obtain scientific and technical advice from experts nominated by the EU Member States, industry organisations and non-governmental organisations. It is expected that a legislative proposal, based on the advice of the Commission Working Groups, will be drafted by the Summer of 2002, after which it will be considered for adoption by the Council and the Parliament, according to the codecision procedure.

1.3 *Directive 86/609/EEC* and ECVAM

In 1986, the Council of Ministers adopted *Directive 86/609/EEC* on the Protection of Laboratory Animals for Experimental and Other Scientific Purposes (Anon, 1986). Article 23 of *Directive 86/609/EEC* states that:

> "The Commission and Member States should encourage research into the development and validation of alternative techniques, which could provide the same level of information as that obtained in experiments using animals, but which involve fewer animals or which entail less painful procedures, and shall take such other steps as they consider appropriate to encourage research in this field."

In October 1991, the European Commission responded to this article by means of a Communication to the Council and the Parliament (Anon, 1991), which established the European Centre for the Validation of Alternative Methods (ECVAM), as a unit of the Joint Research Centre. The mission of ECVAM is to play a leading role at the European level in the independent evaluation of the relevance and reliability of tests for specific purposes, through research on advanced methods and new test development and validation.

1.4 Alternative methods and their application

In the context of laboratory animal use, "alternatives" include *all procedures which can completely replace the need for animal experiments, reduce the number of animals required, or diminish the amount of distress or pain suffered by animals in meeting the essential needs of man and other animals* (Smyth, 1978).

This definition embodies the Three Rs concept proposed by Russell & Burch in *The Principles of Humane Experimental Technique* (1959), and the laws of many countries and *Directive 86/609/EEC* of the European Union now specifically require that *replacement alternatives*, *reduction alternatives* and *refinement alternatives* should be used wherever and whenever possible in biomedical research, testing and education.

The certainty that some additional testing will be necessary in relation to the Future Chemicals Policy, albeit on a scale which cannot be predicted at present, will necessitate the vigorous and dedicated application of all Three Rs, so that any suffering caused to the animals that will have to be used will truly be unavoidable "in meeting the essential needs of man and other animals".

This ECVAM report will primarily focus on replacement alternative tests and testing strategies, but there is a no less urgent need for attention to be focused on the implications of the Future Chemicals Policy in terms of the other two Rs.

There is undoubtedly scope for reducing the numbers of animals required through a reconsideration of group sizes and of test designs (Festing *et al.*, 1998). The statistical method and experimental design adopted for a given study should maximise the scientific output from the use of a given number of animals. Furthermore, the extent and frequency of unavoidable animal suffering could undoubtedly be decreased through the development and use of the earliest and most humane endpoints that meet the scientific objectives (Hendriksen & Morton, 1999). This will be particularly important for the assessment of endpoints for which non-animal replacements are unlikely to be available in the near future (e.g. in chronic toxicity testing).

In addition, there should be an authoritative and independent review of the relevance and reliability of all the animal test procedures that will eventually be required for chemicals for which further testing must be conducted, in compliance with the REACH system.

Furthermore, the Three Rs should not be seen as separate entities, but as complementary elements in a common approach to humane science. For example, it is often more appropriate to think, not of replacements *per se*, but of *non-animal tests and testing strategies*, especially where their use can lead, not to total replacement, but to a reduction in numbers of animal required, through the better prediction of starting doses or the use of fewer treatment groups. Similarly, information obtained in a non-animal test can frequently eliminate the risk that severe effects will be caused to animals necessarily used; for example, to confirm a negative result or a low-level effect.

Thus, the application of an integrated approach to essential testing should involve the simultaneous consideration of all the Three Rs, as well as the integrated (and non-competitive) use of the variety of non-animal tests and non-animal test data which are, or will increasingly become, available. These include:

1. Maximising the use of existing information, including the reasons for producing a chemical and its uses, as well as knowledge of its toxic hazard potential.
2. The use of data concerning the physicochemical properties of chemicals (e.g. stability, solubility, pH, octanol/water partition coefficient, protein binding).
3. Predictions based on structure-activity relationships, including qualitative and quantitative mathematical models, and the use of read-across data from related chemicals.
4. The biokinetic modelling of physiological, pharmacological and toxicological processes.
5. Experiments on lower organisms not classed as "protected animals" (bacteria, fungi, plants, invertebrate animals)
6. Vertebrates at early stages of development (before they become "protected" animals)
7. Studies on *in vitro* systems of various kinds (including whole perfused organs, tissue slices, cell, tissue and organotypic cultures, and subcellular fractions).

8. Human studies (including estimations of occupational and environmental exposure, epidemiological investigations, post-marketing surveillance for medicines, cosmetics and household and agricultural products, and the ethical use and properly-controlled use of human volunteers).

1.5   The role of ECVAM in the formulation of the EU Chemicals Policy

ECVAM has played an advisory role in the formulation of the EU Chemicals Policy, by contributing to the inter-service discussions within the Commission that aim to elaborate the details of the new policy. To help ECVAM in this role, ECVAM established a Working Group on Chemicals (referred to hereafter as the Working Group) in July 2001. The remit given to the Working Group was to propose, by the end of 2001, a strategy on alternative (non-animal) methods in relation to the emerging Chemicals Policy. Details of the membership of the Working Group are given in Table 1.3

At its first meeting on 24-25 July 2001, the Working Group decided to focus its discussions and recommendations on the use of alternative methods for the hazard assessment of toxicological endpoints for human health (rather than ecotoxicological endpoints), with particular emphasis on the toxicological endpoints currently required in the EU for chemicals testing, i.e. the endpoints defined in Annexes VII and VIII of *Directive 67/548/EEC* (Tables 1.4-1.6).

The Working Group decided to produce a detailed review of the current status of alternative methods for chemicals testing, which would also include proposals for the strategic use of alternative tests, and a three-stage action plan, containing proposals reflecting the short-, medium- and long-term prospects for the development and validation of alternative tests (Table 1.7).

To help with the production of the review document, the Working Group recommended that nine ECVAM Focus Groups should be established, which would focus their efforts on nine areas corresponding to the different toxicological endpoints. The membership of the ECVAM Focus Groups on Toxicological Endpoints is given in Table 1.8. Additional experts were also consulted on specific issues, when necessary.

The Working Group had further meetings on 1-2 October, 2001, and on 21-22 January 2002, to discuss the progress being made with the ECVAM review document, and to exchange ideas for the strategic use of alternative tests. On 23 January, a draft version of the document was discussed at ECVAM with a number of experts and stakeholders who had not previoously been involved in its production (Table 1.9).

In addition to the White Paper (Anon, 2001), the Working Group based its discussions and recommendations on a number of other documents, including:

1. the conclusions on the chemicals policy reached by the European Environment Council, in its meeting of 7 June 2001 (Anon, 2001);
2. a document published by the British Union for the Abolition of Vivisection (BUAV, 2001), containing a proposal for a non-animal testing strategy;

3. documents published by the Medical Research Council's Institute for Environment and Health, UK, consisting of assessments of: a) the implications of the future chemicals policy, in terms of financial cost and animal use (IEH, 2001a); and b) the feasibility of replacing current regulatory *in vivo* tests *with in vitro* tests (IEH, 2001b); and

4. a document about the TestSmart Program in the USA (Green *et al.*, 2001). Further information on this programme is available from: http://caat.jhsph.edu/programs/workshops/testsmart/hpv-intro.htm.

1.6   The role of ECVAM in the implementation of the EU Chemicals Policy

ECVAM will also play a role in the implementation of the future chemicals policy. In Section 3.2 of the White Paper, it is stated that:

> "One of the major tasks of the European Centre for the Validation of Alternative Methods (ECVAM) of the Joint Research Centre of the Commission is to validate alternative methods that reduce, refine or replace animal experiments …
>
> ECVAM's central role will be maintained, and it is expected that the development of alternative methods will be accelerated. Further research will be carried out both at Community and national level, in order to develop and validate novel testing strategies involving fewer or no animals, while enhancing the relevant information that can be obtained from testing without simultaneously increasing the number of animals involved."

An additional role for ECVAM is referred to in Section 6 of the White Paper, which outlines the provisions for the testing and evaluation of existing substances, and envisages the establishment of an advisory task force of experts from the Member States. These experts will be assigned various responsibilities, including (for substances exceeding 1000 t.p.a.) the proposal of additional testing programmes. It is clearly stated that these proposals should be made "in cooperation with ECVAM."

The future role of ECVAM is also referred to in the conclusions of the European Environment Council on the chemicals policy (Anon, 2001), reached during its 2355th session, held on 7 June 2001. In paragraph 23 of the Council conclusions, it is stated that:

> "Adequate resources must be provided for research, development and validation of globally accepted test guidelines for alternative *in vitro* test methods, so that work can be accelerated at all levels. Activities under the new Framework Programme for Research should consider these requirements among its priorities. In addition to promoting this issue in ECVAM (European Centre for the Validation of Alternative Methods), the Community should play a more active role in the OECD, to encourage wider adoption of validated, alternative, non-animal testing methods."

In addition to its role in the validation of alternative methods, and in promoting the regulatory acceptance of alternative methods (Worth & Balls, 2001), ECVAM also

plays an important role as an information centre and in providing a forum for discussion. For the dissemination of information, ECVAM is developing a Scientific Information Service (ECVAM SIS), which is available on the internet (http://ecvam-sis.jrc.it. This contains comprehensive information about validation studies, including the experimental protocols of validated tests and tests under development, and test data generated in validation and other studies. ECVAM provides a forum for scientific discussion through its workshops, at which experts from academia, industry and other communities discuss the state-of-the-art in a particular field, and make recommendations for further progress. To date, the reports of more than 45 ECVAM workshops have been published (Table 1.10).

Table 1.1. The proposed schedule for the testing of 30, 100 existing substances (Anon, 2001a)

Number of substances	Volume (tonnes per annum)	Deadline for testing
2600	> 1,000	end of 2005
2900	100-1000	end of 2008
4,600	10-100	end of 2012
20,000	1-10	end of 2012

Table 1.2. The dependence of testing requirements on tonnage (Anon, 2001a)

Volume (tonnes per annum)	Testing required
1-10 t	"Testing should generally be limited to *in vitro* methods"
10-100 t	Base-set testing
100-1000 t	Base-set testing + Level 1 testing, i.e. "substance-tailored testing for long-term effects"
> 1000 t	Base-set testing + Level 2 testing, i.e. "additional substance-tailored testing for long-term effects"

Table 1.3. Membership of the ECVAM Working Group on Chemicals.

Member	Contact details
Professor Michael Balls - Chairman	ECVAM, IHCP, JRC, Italy Fax: +39 0332 786297 E-mail: michael.balls@jrc.it E-mail from 01.07.02: michael.balls@frame.org.uk
Dr Andrew Worth – Secretary	ECVAM, IHCP, JRC, Italy Fax: +39 0332 785845 E-mail: andrew.worth@jrc.it
Mr Martin Barratt	Marlin Consultancy, Bedford, UK Fax: + 44 1234 721 533 E-mail: martin.d.barratt@btinternet.com
Dr Bas Blaauboer	Utrecht University, Utrecht, The Netherlands Fax: +31 30 253 5077 E-mail: b.blaauboer@iras.uu.nl
Dr Phil Botham	Syngenta, Macclesfield, UK Fax: +44 1625 590 249 E-mail: phil.botham@syngenta.com
Professor Robert Combes	Fund for the Replacement of Animals in Medical Experiments (FRAME), Nottingham, UK Fax: + 44 115 950 3570 E-mail: bob@frame.org.uk
Professor Johannes Doehmer	GenPharmTox BioTech AG, Munich, Germany Fax: + 49 89 289 14 183 E-mail: johannes.doehmer@genpharmtox.de

Dr Julia Fentem	Safety and Environmental Assurance Centre, Unilever, Sharnbrook, UK Fax: + 44 1234 264722 E-mail: julia.fentem@unilever.com
Dr Manfred Liebsch	ZEBET, BgVV, Berlin, Germany Fax: + 49 1888 412 2958 E-mail: liebsch.zebet@bgvv.de
Professor Horst Spielmann	ZEBET, BgVV, Berlin, Germany Fax: + 49 1888 412 2958 E-mail: spielmann.zebet@bgvv.de

Table 1.4. Base-set Testing Requirements for Human Health Endpoints.[*]

Endpoint	EU test methods
Acute toxicity	B.1: acute toxicity (oral) – DELETED ON 25/1/01 B.1bis: acute toxicity (oral) fixed dose method B.1tris: acute toxicity (oral) – acute toxic class method B.2: acute toxicity (inhalation) B.3: acute toxicity (dermal)
Irritation	B.4: acute toxicity (skin irritation) B.5: acute toxicity (eye irritation)
Corrosivity	B.40: skin corrosion
Skin and respiratory sensitisation	B.6: skin sensitisation No Annex V method for respiratory sensitisation
Repeated dose toxicity	B.7: repeated dose (28 days) toxicity (oral) B.8: repeated dose (28 days) toxicity (inhalation) B.9: repeated dose (28 days) toxicity (dermal)
Mutagenicity and genotoxicity	B.10: mutagenicity (*in vitro* mammalian chromosome aberration test) B11: mutagenicity (*in vivo* mammalian bone-marrow chromosome aberration test) B.12: mutagenicity mammalian erythrocyte micronucleus test B.13/14: mutagenicity – reverse mutation test using bacteria B.15: gene mutation – *saccharomyces cerevisae* B.16: mitotic recombination - *saccharomyces cerevisae* B.17: mutagenicity - *in vitro* mammalian cell gene mutation test B.18: DNA damage and repair – unscheduled DNA synthesis – mammalian cells *in vitro* B.19: sister chromatid exchange assay *in vitro* B.20: sex-linked recessive lethal test in *drosophila melanogaster* B.21: *in vitro* mammalian cell transformation test B.22: rodent dominant lethal test B.23: mammalian spermatogonial chromosome aberration test B.24: mouse spot test B.25: mouse heritable translocation B.39: unscheduled DNA synthesis (UDS) test with mammalian liver cells *in vivo*

[*]Based on Annexes VII A, B and C of *Directive 67/548/EEC*

Table 1.5. Level 1 Testing Requirements for Human Health Endpoints.[*]

Endpoint	EU test methods
Sub-chronic and/or chronic toxicity	B.26: sub-chronic oral toxicity test: 90-day repeated oral dose study using rodent species B.27: sub-chronic oral toxicity test: 90-day repeated oral dose study using non-rodent species B.28: sub-chronic dermal toxicity test: 90-day repeated dermal dose study using rodent species B.29: sub-chronic inhalation toxicity test: 90-day repeated inhalation dose study using rodent species B.30: chronic toxicity test
Developmental toxicity	B.31: teratogenicity test – rodent and non-rodent
Fertility study	B.34: one-generation reproduction toxicity test B.35: two-generation reproduction toxicity test
Additional mutagenicity studies	
Toxicokinetics	B.36: toxicokinetics

[*]Based on Annex VIII of *Directive 67/548/EEC*

Table 1.6. Level 2 Testing Requirements for Human Health Endpoints.[*]

Endpoint	EU test method
Chronic toxicity	B.30: chronic toxicity test
Carcinogenicity	B.21: *in vitro* mammalian cell transformation test B.32: carcinogenicity test B.33: combined chronic toxicity/carcinogenicity test
Developmental toxicity	Using species not used in Level 1 study
Developmental toxicity	For peri-natal and post-natal effects
Fertility study	Extended B.35: three-generation reproduction toxicity test
Additional pharmacokinetic studies	To cover, for example, biotransformation
Additional organ or system toxicity	B.7 includes neurotoxicity and immunotoxicity B.37: delayed neurotoxicity of organophosphorus substances following acute exposure B.38: delayed neurotoxicity of organophosphorus substances 28 day repeated dose study

Footnote

[*]Based on Annex VIII of *Directive 67/548/EEC*

Table 1.7. A proposed time-frame for the development and validation of alternative methods for chemicals testing.

Priority	Time-frame for completion of activities
1) Short-term	2003
2) Medium-term	2006
3) Long-term	2010

Footnote

Adherence to this time-frame will depend on many factors, including the availability of human and financial resources.

Table 1.8. Membership of the ECVAM Focus Groups on Toxicological Endpoints.

ECVAM Working Group	ECVAM Responsible	External Partners
1. Acute systemic toxicity	Silvia Casati	Manfred Liebsch (ZEBET) Richard Clothier (FRAME)
2. Acute local toxicity a) Skin irritation / corrosion b) Eye irritation / corrosion	 Andrew Worth Valérie Zuang	 Julia Fentem (Unilever) Philip Botham (Syngenta)
3. Sensitisation (skin and respiratory)	Valérie Zuang	David Basketter (Unilever) Martin Barratt (Marlin Consultancy)
4. Repeat-dose toxicity	Pilar Prieto	Walter Pfaller (University of Innsbruck)
5. Genotoxicity and carcinogencity	Enrico Sabbioni	Robert Combes (FRAME)

6.	**Reproductive toxicity**	Susanne Bremer	Horst Spielmann (ZEBET)
7.	**Neurotoxicity**	Sandra Coecke	Erik Walum (Biovitrum)
8.	**Toxicokinetics – absorption and distribution**		
a)	*in vitro* barrier function tests (blood-brain barrier, gastrointestinal tract, kidney)	Pilar Prieto	Per Garberg (Biovitrum)
b)	computer modelling of absorption and distribution	Andrew Worth	Bas Blaauboer (RITOX) Mark Cronin (Liverpool John Moores University)
9)	**Toxicokinetics – metabolism**	Sandra Coecke	Johannes Doehmer (GenPharmTox BioTech AG)

Table 1.9. Participants at the ECVAM Stakeholders Meeting on 23 January 2002.[*]

Bernward Garthoff	Bayer AG, Germany E-mail: bernward.garthoff.bg@bayer-ag.de
Paul Harrison	Institute for Environment and Health, Leicester University, UK E-mail: ptch1@leicester.ac.uk
Ian Indans	Health & Safety Executive, UK E-mail: Ian.Indans@hse.gsi.gov.uk
Mike Jackson	Consultant, Nuthampstead, Herts, UK E-mail: mjackson@sakernet.com
Gill Langley	British Union for the Abolition of Vivisection, UK E-mail: gill@drhadwentrust.org.uk
Beatrice Lucaroni	DG RTD/E.4, European Commission E-mail: beatrice.lucaroni@cec.eu.int
Barry Phillips	Royal Society for the Prevention of Cruelty to Animals, UK E-mail: bphillips@rspca.org.uk
Juan Riego-Sintes	ECB, IHCP, JRC, European Commission E-mail: juan.riego-sintes@jrc.it
Arturo Sanabria	DG SANCO/C.2, European Commission E-mail: arturo.sanabria-tienza@cec.eu.int
Regina Schumann	DG ENTR/F.3, European Commission E-mail: regina.schumann@cec.eu.int
Kees van Leeuwen	Director of IHCP, JRC, European Commission E-mail: kees.van-leeuwen@cec.eu.int
Ursula Sauer	Akademie für Tierschutz, Neubiberg, Germany E-mail: ursula.sauer@tierschutzakademie.de

[*]Excluding members of ECVAM staff and the Working Group on Chemicals

Table 1.10. List of ECVAM Workshops.

Number	Title	Venue / date	Publication details
1	The Practical Applicability of Hepatocyte Cultures in Routine Testing	Angera, Italy 19-23 October 1993	Blaauboer *et al.* (1994) *ATLA* 22, 231-241
2	*In Vitro* Phototoxicity Testing	Angera, Italy 13-17 December 1993	Spielmann *et al.* (1994) *ATLA* **22**, 314-348
3	*In Vitro* Neurotoxicity Testing	Angera, Italy 7-11 February 1994	Atterwill *et al.* (1994) *ATLA* **22**, 350-362
4	Alternatives to Animal Testing in the Quality Control of Immunobiologicals: Current Status and Future Prospects	Utrecht, The Netherlands 16-17 April 1994	Hendriksen *et al.* (1994) *ATLA* **22**, 420-434
5	Practical Aspects of the Validation of Toxicity Test Procedures	Amden, Switzerland 24-28 January 1994	Balls *et al.* (1995) *ATLA* **23**, 129-147
6	A Prevalidation Study on *In Vitro* Skin Corrosivity Testing	Angera, Italy 12-14 January 1994	Botham *et al.* (1995) *ATLA* **23**, 219-255
7	Development and Validation of Non-animal Tests and Testing Strategies: the Identification of a Coordinated Response to the Challenge and the Opportunity Presented by the Sixth Amendment to the Cosmetics Directive (*76/768/EEC*)	Angera, Italy 11-13 April 1994	Balls *et al.* (1995) *ATLA* **23**, 398-409
8	The Integrated Use of Alternative Approaches for Predicting Toxic Hazard	Angera, Italy 23-27 January 1995	Barratt *et al.* (1995) *ATLA* **23**, 410-429
9	Safety and Efficacy Testing of Hormones and Related Products	Düsseldorf, Germany 25-27 November 1994	Garthoff *et al.* (1995) *ATLA* **23**, 699-712
10	Nephrotoxicity Testing *In Vitro*	Angera, Italy 16-20 May 1994	Hawksworth *et al.* (1995) *ATLA* **23**, 713-727
11	The Three Rs: The Way Forward	Sheringham, UK 30 May-3 June 1996	Balls *et al.* (1995) *ATLA* **23**, 838-866
12	Screening Chemicals for Reproductive Toxicity: the Current Alternatives	Angera, Italy 22-26 February 1994	Brown *et al.* (1995) *ATLA* **23**, 868-882
13	Methods for Assessing Percutaneous Absorption	Angera, Italy 30 May-3 June 1994	Howes *et al.* (1996) *ATLA* **24**, 81-106
14	The Use of *In Vitro* Systems for Evaluating Haematotoxicity	Angera, Italy 29 May-2 June 1995	Gribaldo *et al.* (1996) *ATLA* **24**, 211-231
15	The Use of Biokinetics and *In Vitro* Methods in Toxicological Risk Evaluation	Utrecht, The Netherlands 21-23 March 1995	Blaauboer *et al.* (1996) *ATLA* **24**, 473-497
16	Acute Toxicity Testing *In Vitro* and the Classification and Labelling of Chemicals	Angera, Italy 18-22 April 1994	Seibert *et al.* (1996) *ATLA* **24**, 499-510
17	Alternatives to the Animal Testing of Medical Devices	Copenhagen, Denmark 24-26 November 1995	Svendsen *et al.* (1996) *ATLA* **24**, 659-669
18	*In Vitro* Tests for Respiratory Toxicity	Angera, Italy 14-18 November 1994	Lambré *et al.* (1996) *ATLA* **24**, 671-681
19	Alternative Methods for Skin Sensitisation Testing	Angera, Italy 24-28 April 1995	de Silva *et al.* (1996) *ATLA* **24**, 683-705
20	The Use of Tissue Slices for Pharmacotoxicology Studies	Angera, Italy 27-31 May 1996	Bach *et al.* (1996) *ATLA* 24, 893-923

Number	Title	Venue / date	Publication details
21	The Production of Avian (Egg Yolk) Antibodies: IgY	Berlin, Germany 22-24 March 1996	Schade *et al.* (1996) *ATLA* **24**, 925-934
22	Pharmacokinetics in Early Drug Research	Bath, UK 27-29 March 1996	Leahy *et al.* (1997) *ATLA* **25**, 17-31
23	Monoclonal Antibody Production. ECVAM Workshop Report 23	Angera, Italy 19-22 November 1996	Marx *et al.* (1997) *ATLA* **25**, 121-137
24	The Development and Validation of Expert Systems for Predicting Toxicity	Angera, Italy 1-4 October 1996	Dearden *et al.* (1997) *ATLA* 25, 223-252
25	Current Status and Future Developments of Databases on Alternative Methods	Neubiberg, Germany 12-15 September 1996	Janusch *et al.* (1997) *ATLA* 25, 411-422
26	Genetically Engineered Cell Lines: Characterisation and Applications in Toxicity Testing	Angera, Italy 26-27 February 1996	Wiebel *et al.* (1997) *ATLA* **25**, 625-639
27	Issues Relating to the Release of Proprietary Information and Data for Use in the Validation of Alternative Methods	Munich, Germany 9-11 May 1997	Todd *et al.* (1998) *ATLA* **26**, 13-20
28	The Use of Transgenic Animals in the European Union	Southwell, UK 7-11 April 1997	Mepham *et al.* (1998) *ATLA* **26**, 21-43
29	Reducing the Use of Laboratory Animals in Biomedical Research: Problems and Possible Solutions	Southwell, UK 12-15 January 1998	Festing *et al.* (1998) *ATLA* **26**, 283-301
30	Non-animal Tests for Evaluating the Toxicity of Solid Xenobiotics	Angera, Italy 28-31 October 1997	Fubini *et al.* (1998) *ATLA* **26**, 579-617
31	Validation of Alternative Methods for the Potency Testing of Vaccines	Angera, Italy 14-16 November 1997	Hendriksen *et al.* (1998) *ATLA* **26**, 747-761
32	The Availability of Human Tissue for Biomedical Research	Barnsdale, Rutland, UK 18-22 May 1998	Anderson *et al.* (1998) *ATLA* **26**, 763-777
33	Alternatives to the Use of Animals in Higher Education	Crete, Greece 8-10 May 1998	Van der Valk *et al.* (1998) *ATLA* **27**, 39-52
34	Eye Irritation Testing: The Way Forward	Egham, UK 15-17 June 1998	Balls *et al.* (1999) *ATLA* **27**, 53-77
35	The Production of Polyclonal Antibodies in Laboratory Animals	Utrecht, The Netherlands 20-22 March 1998	Leenaars *et al.* (1999) *ATLA* **27**, 79-102.
36	The Potential Use of Non-invasive Methods in the Safety Assessment of Cosmetic Products	Brussels, Belgium 10-12 March 1998	Rogiers *et al.* (1999) *ATLA* **27**, 515-537
37	The Principles of Good Laboratory Practice: Application to *In Vitro* Toxicology Studies	Angera, Italy 6-9 December 1998	Cooper-Hannan *et al.* (1999) *ATLA* **27**, 539-577
38	The Use of Human Keratinocytes and Human Skin Models for Predicting Skin Irritation	Utrecht, The Netherlands 9-11 November 1997	Van de Sandt *et al.* (1999) *ATLA* **27**, 723-743
39	Cell Transformation Assays as Predictors of Human Carcinogenicity	Angera, Italy 12-16 October 1998	Combes *et al.* (1999) *ATLA* **27**, 745-767
40	Biomarkers as Predictive Tools in Toxicity Testing	Burnham Market, UK 5-8 October 1998	Benford *et al.* (2000) *ATLA* **28**, 119-131

Number	Title	Venue / date	Publication details
41	Three Rs Approaches in the Production and Quality Control of Avian Vaccines	Langen, Germany 11-13 June 1999	Bruckner *et al.* (2000) *ATLA* **28**, 241-258
42	The Second ECVAM Workshop on Phototoxicity Testing	Berlin, Germany 22-27 June 1999	Spielmann *et al.* (2000) *ATLA* **28**, 777-814
43	Novel Pyrogen Tests Based on the Human Fever Reaction	Konstanz, Germany 16-20 January 2000	Hartung *et al.* (2001) *ATLA* **29**, 99-123
44	The Establishement of Human Research Tissue Banking in the UK and Several Western European Countries	Birmingham, UK 8-10 September 2000	Anderson *et al.* (2001) *ATLA* **29**, 125-134
45	Novel Advanced *In Vitro* Methods for Long-term Toxicity Testing	Innsbruck, Austria 17-21 May 1999	Pfaller *et al.* (2001) *ATLA* **29**, 393-426
46	*In Vitro* Models of the Intestinal Barrier	Paris, France 4-5 March 1999	Le Ferrec *et al.* (2001) *ATLA* **29**, 649-668

2.   THE PRINCIPLES AND PROCEDURES OF VALIDATION

2.1 Introduction

The aim of this chapter is to give a brief overview of the principles of validation, and of the manner in which they are applied by ECVAM, with reference to the main publications in the field. A short review by Balls & Fentem (1999) explains the key concepts of validation and regulatory acceptance, whereas a consolidated account based on the many papers that describe ECVAM's principles and procedures is in preparation.

2.2 Alternative methods, prediction models and validation

An alternative method is any method that can be used to replace, reduce or refine the use of animal experiments in biomedical research, testing, or education. The "alternatives" concept is attributed to Russell & Burch (1959), who defined three types of alternatives: *reduction alternatives*, which obtain a comparable level of information from the use of fewer animals, or more information from the same number of animals; *refinement alternatives*, which alleviate or minimise potential pain, suffering and distress; and *replacement alternatives*, which permit a given purpose to be achieved without using animals. The three types of alternative procedure are referred to collectively as the "Three Rs".

The prediction model (PM) plays an important role in the validation process. As described by Archer *et al.* (1997), an alternative method for the replacement (or partial replacement) of an animal test can be thought of as the combination of a test system and a PM. The test system provides a means of generating physicochemical or *in vitro* data for chemicals of interest, whereas the PM is an unambiguous algorithm for converting these data into predictions of a pharmacotoxicological endpoint in animals or humans. The role of the PM in validation was discussed earlier by Bruner *et al.* (1996a), who also defined criteria for the adequacy of PMs, and described the use of computer simulations, based on the PM concept, for judging the performance of alternative tests. Subsequently, there were a number of editorials in *ATLA*, which debated the use of PMs in validation studies (Bruner *et al.*, 1996b; Archer, 1996; Holzhütter & Spielmann, 1996). A consensus view on the use of PMs was described in a joint report by the ECVAM Task Forces on validation and on biostatistics (Archer *et al.*, 1997). More recently, a consolidated description of the use of PMs, including their incorporation into tiered testing strategies, has been published from ECVAM (Worth & Balls, 2001a).

The validation of an alternative method is the process by which the relevance and reliability of the method are established for a particular purpose (Balls *et al.*, 1990). In the context of a replacement *test* method, relevance refers to the scientific basis of the test system, and to the predictive capacity of an associated PM, whereas reliability refers to the reproducibility of test results, both within and between laboratories, and over time (Figure 2.1). The "purpose" of an alternative method refers to its intended application, such as the regulatory testing of chemicals for a specific toxicological endpoint, such as skin corrosivity. In other words, to validate (i.e. to establish the

scientific validity of) an alternative test, it is necessary to demonstrate that for its stated purpose:

1) the test system has a sound scientific basis;
2) the predictions made by the PM are sufficiently accurate; and
3) the results generated by the test system are sufficiently reproducible within and between laboratories, and over time.

These conditions can be referred to as the criteria for scientific relevance, predictive relevance, and reliability, respectively.

2.3 The evolution of alternative methods and the ECVAM validation process

Five main stages in the evolution of new test methods have been identified (Balls *et al.*, 1995; Fentem & Balls, 1997):

1) test development
2) prevalidation
3) validation
4) independent assessment
5) progression toward regulatory acceptance

The main role of ECVAM is to coordinate, at the European Union level, the validation of alternative methods, although ECVAM also plays a role in promoting test development and encouraging the progression of scientifically valid methods toward regulatory acceptance. The roles of ECVAM, and of the ECVAM Scientific Advisory Committee (ESAC) have been described (Worth & Balls, 2001b).

The ECVAM validation process consists of two main stages: prevalidation and formal validation.

A prevalidation study is a small-scale inter-laboratory study designed to refine the protocol and PM of a test method, and to obtain a preliminary assessment of its relevance and reliability. The prevalidation process is divided into three consecutive phases (Curren *et al.*, 1995): method refinement (phase I), method transfer (phase II), and method performance (phase III). During phase I, the protocol and PM of a test method are refined in a single laboratory (with prior experience in the use of the test), whereas during phase II, an assessment is made of the transfer ability of the method to a second laboratory, making any necessary refinements to the protocol and PM. During phase III, the relevance and reliability of the test are assessed under blind conditions in three or more laboratories (which generally include the first two laboratories). An important outcome from a prevalidation study is that an optimised protocol is identified that could be used in a formal validation study.

A formal validation study is a larger-scale inter-laboratory study, performed under blind conditions, and designed to obtain a more definitive assessment of relevance and reliability. A formal validation study can be thought of as a larger-scale version of the phase III stage of prevalidation, in which a larger number of chemicals are tested (although not necessarily in a larger number of laboratories). In general, newly-developed test methods enter the prevalidation process, and following successful

prevalidation, proceed to formal validation. In some cases, however, a new test may be sufficiently "similar" (in terms of its structural characteristics and reliability) to another test, the scientific validity of which has already been established, for a prevalidation study to be sufficient to establish the validity of the new method (Balls, 1997). This process, called "catch-up" validation, has been applied to the EpiDerm^{TM} human skin model for skin corrosivity (Liebsch *et al.*, 2000), following the successful formal validation of the EPISKIN^{TM} human skin model (Fentem *et al.*, 1998). A schematic representation of the ECVAM validation process is given in Figure 2.2.

ECVAM's criteria for development and validation have been defined in a number of documents, and are summarised in Section 2.4 below (Balls & Karcher, 1995; Balls *et al.*, 1995a; Fentem & Balls, 1997; Balls & Fentem, 1997). There are thought to be no major differences between these criteria and the criteria subsequently defined by ICCVAM (NIH, 1997; NIH 1999) and the OECD (OECD, 1996).

Following the completion of an ECVAM prevalidation or validation study, a report on the outcome of the study is submitted to ECVAM, so that it can be communicated to, and considered by, the ESAC. If the ESAC is satisfied that ECVAM's criteria for test development and validation have been satisfied, an ESAC statement endorsing the scientific validity of the method is issued.

The ESAC also considers the scientific validity of methods that have been assessed by a weight-of-evidence consideration of existing information (retrospective validation), rather than by the conduct of new practical work in prevalidation and/or validation studies (prospective validation). The ESAC endorsement process, and the subsequent steps by which validated methods are submitted for regulatory consideration at the EU and OECD levels, have been described in detail elsewhere (Worth & Balls, 2001b).

2.4 ECVAM's criteria for test development and validation

For a method to be considered ready to enter the validation process, the following should be available, to establish that is has been sufficiently well developed:

1. A definition of the scientific purpose of the method, and of its proposed practical application.
2. A description of the basis of the method.
3. The case for its relevance. In the case of a test method for the prediction of an *in vivo* pharmacotoxicological endpoint, this should refer to the mechanistic relevance of the test, and to any preliminary evidence supporting the predictive capacity of the test.
4. An explanation of the need for the method in relation to existing *in vivo* (animal or human) methods (refer to relevant test guidelines and legislation) and other non-animal methods.
5. An optimised protocol, including: a) any necessary standard operation procedures; b) a specification of endpoints and endpoint measurements; c) the method for deriving and expressing results; d) the interpretation of the results in terms of one or more *in vivo* pharmacotoxicological endpoints, by means of a prediction model; and e) the use of adequate controls.
6. A statement about the limitations of the test.

7. Evidence of intra-laboratory reproducibility, and if available, inter-laboratory transferability.

In an independent assessment of the outcome a prevalidation or validation study, the following (validation criteria) should be considered:

1. Clarity of defined goals.
2. Quality of overall design.
3. Independence of management.
4. Independence of selection, coding and distribution of test materials.
5. Independence of data collection and analysis.
6. Number and properties of test materials.
7. Quality and interpretation of results
8. Performance of the method(s) in relation to the predetermined goals of the study.
9. Reporting of outcome in the peer-review literature.
10. Availability of raw data.
11. Independence of assessment of outcome.

2.5 Practical and logistical aspects of prevalidation and validation studies

The design, management and conduct of prevalidation and validation have been described in a number of publications. Curren *et al.* (1995) focused on the role of prevalidation, describing in detail the three phases of this process. Balls *et al.* (1995) described: a) the role of the Management Team in a validation study; b) the selection of alternative tests, laboratories and chemicals; c) compliance with safety standards and the principles of Good Laboratory Practice (GLP); d) the collection and analysis of experimental data; and d) the assessment of the outcome of the validation studies. Problems encountered in some early validation studies, and difficulties associated with the selection and use of *in vivo* data, and the comparison of *in vitro* and *in vivo* endpoints, were discussed by Balls & Fentem (1997). Recommendations concerning the application of GLP principles to *in vitro* toxicology, and in particular, to validation studies, has been published by Cooper-Hannan *et al.* (1999).

Recommendations for the application of statistical methods in the validation process were made by Holzhütter *et al.* (1996), and there are several examples of the use of biostatistical methods for the development and assessment of PMs (Genshow *et al.*, 2000; Spielmann *et al.*, 1996; Worth & Cronin, 2001)

2.6 Examples of prevalidation and validation studies

Reports on a number of ECVAM prevalidation and validation studies have been published (for example, Botham *et al.*, 1995; Balls *et al.*, 1995b). In terms of their compliance with currently accepted principles and procedures, the ECVAM skin irritation prevalidation study (Fentem *et al.*, 2001) can be regarded as a model for an ECVAM prevalidation study, whereas the ECVAM skin corrosivity validation study (Fentem *et al.*, 1998) and the EU/COLIPA phototoxicity validation study (Spielmann *et al.*, 1998) can be regarded as model validation studies. The selection of test chemicals for a validation study is well illustrated by Barratt *et al.* (1998).

As an example, the ECVAM skin corrosivity validation study (Barratt *et al.*, 1998; Fentem *et al.*, 1998) was conducted as a follow-up to an ECVAM prevalidation study (Botham *et al.*, 1995). The main objectives of the validation study were to: a) identify tests capable of discriminating corrosives from non-corrosives for selected types of chemicals and/or all chemicals; and b) determine whether these tests could identify correctly known R35 and R34 chemicals. The tests evaluated were the rat skin transcutaneous electrical resistance (TER) assay, CORROSITEX^{TM}, the Skin^{2TM} ZK1350 corrosivity test and EPISKIN^{TM}. Each test was conducted in three independent laboratories (to assess interlaboratory transferability), on at least two occasions (to assess interlaboratory reproducibility), with sixty coded chemicals. Since three laboratories were used to perform experimental work on each of four tests, a total of 12 laboratories were involved. These laboratories were based in both the EU and the USA, which illustrates the fact validation studies are international collaborative exercises.

All the tests evaluated showed acceptable intralaboratory and interlaboratory reproducibilities, and the TER, Skin^2 and EPISKIN tests proved applicable to testing a diverse group of chemicals of different physical forms, including organic acids, organic bases, neutral organics, inorganic acids, inorganic bases, inorganic salts, electrophiles, phenols and soaps/surfactants. Two of the four tests evaluated, the TER assay and EPISKIN, met the criteria agreed by the Management Team concerning acceptable underprediction and overprediction rates, so they could be considered scientifically validated for use as replacements for the animal test for distinguishing between corrosive and non-corrosive chemicals for all of the chemical types studied EPISKIN was the only test able to distinguish between known R35 and R34 chemicals, for all of the chemical types included, on an acceptable number of occasions. The corrosive potentials of about 40% of the test chemicals could not be assessed with CORROSITEX, and the assay did not meet the criteria so that it could be considered acceptable as a replacement test. However, CORROSITEX was considered to be valid for testing specific classes of chemicals, such as organic bases and inorganic acids. The Skin^2 assay did not meet the criteria required for it to be considered scientifically validated. On the basis of the successful outcome of this validation study, the ESAC subsequently endorsed the use of the EPISKIN (ECVAM, 1998a) and TER (ECVAM, 1998b) methods.

2.7 The time required for validation and regulatory acceptance

There is widespread dissatisfaction at the rate of adoption of validated methods into regulatory requirements and testing guidelines. As an illustration, Table 2.1 gives the timing of the most important steps leading from the ECVAM prevalidation study on in vitro methods for skin corrosivity, to the validation study and the regulatory acceptance of two skin corrosivity methods at the EU level.

The rate-limiting steps in the evolution of alternative methods appear to be the rate at which suitably developed methods become available for entry into the prevalidation process, and the rate at which scientifically validated methods are accepted by regulatory authorities. The 'ECVAM-dependent' part of the process, which generally comprises both the prevalidation and validation steps, typically takes 4-6 years. However, adoption of the fast-track validation procedure, catch-up validation

(Liebsch *et al.*, 2000; ECVAM, 2001), for appropriate methods, can reduce the time taken for validation to less than two years.

Figure 2.1. A schematic representation of an alternative test and its performance properties.
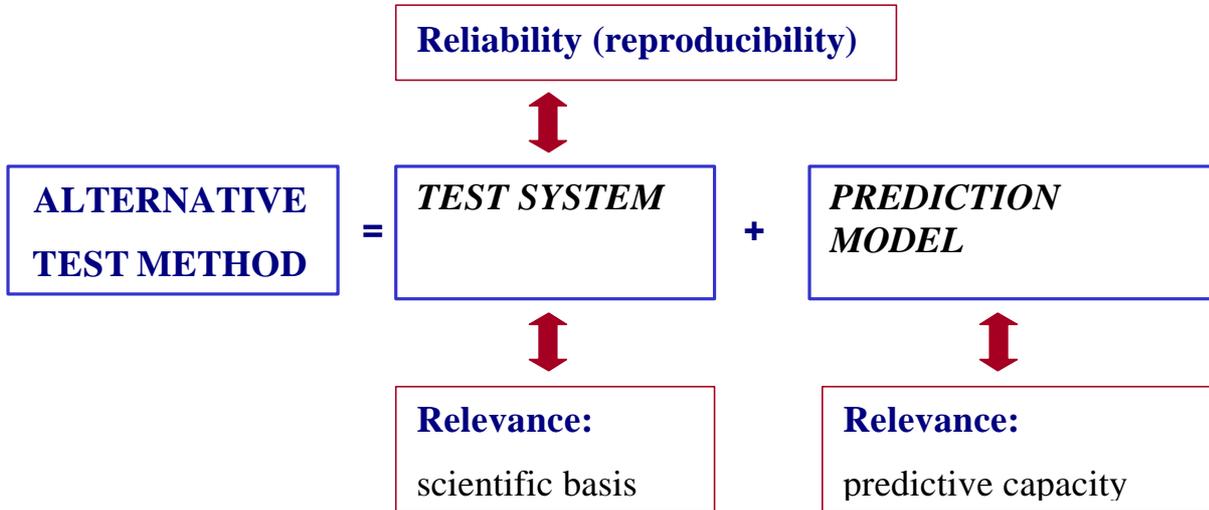
Figure 2.2. A schematic representation of the ECVAM validation process.

```
┌─────────────────────┐                              ┌──────────────────────────┐
│  New test method    │                              │  "Similar" test method   │
└─────────────────────┘                              └──────────────────────────┘
          │                                                       │
          ▼                                                       ▼
┌─────────────────────┐                              ┌──────────────────────────┐
│   Prevalidation     │                              │   Catch-up  validation   │
└─────────────────────┘                              └──────────────────────────┘
          │                                                       ▲
          ▼                                                       │
┌─────────────────────┐          Define structural and           │
│  Formal validation  │───────▶  performance                      │
└─────────────────────┘          characteristics and             │
          │                      criteria                         │
          ▼                                                       │
     ┌─────────────────────────────────┐                         │
     │  Scientifically validated       │◀────────────────────────┘
     │  test method                    │
     └─────────────────────────────────┘
```
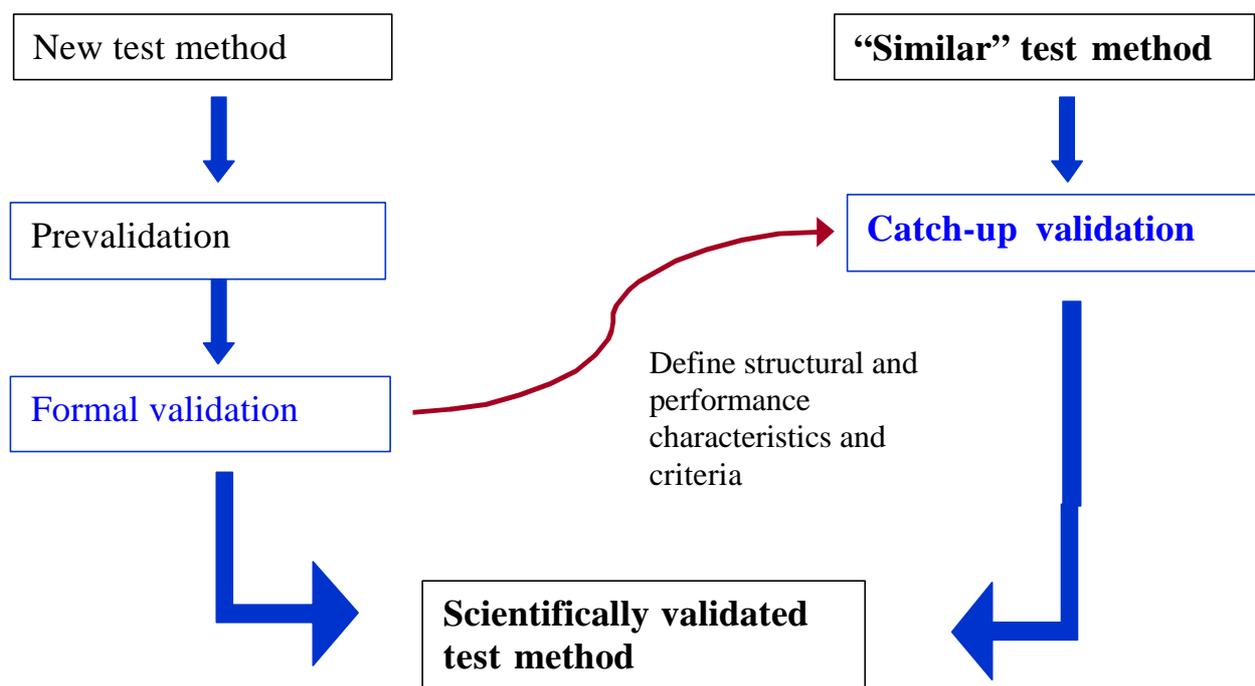
Footnote

In the ECVAM process, a scientifically validated method is one that has been endorsed by the ECVAM Scientific Advisory Committee (ESAC). If the method is appropriate for chemicals testing, a draft Annex V guideline, incorporating the method, will be submitted to the EU Competent Authorities for *Directive 67/548/EEC* for consideration for regulatory acceptance and application.

Table 2.1. Main steps in the validation and regulatory acceptance of *in vitro* methods for skin corrosivity.

Activity	Date
1. Prevalidation study	1993-1994
2. Meeting to discuss proposed validation study	8-9 February 1995
3. Publication of outcome of prevalidation study in *ATLA*	April 1995
4. Chemical selection meeting	19 September 1995
5. Publication of background information on study	November 1995
6. 1st Management Team meeting	11 January 1996
7. Confirmation of participating laboratories	March 1996
8. Meeting to select final set of 60 test chemicals	15 April 1996
9. 2nd Management Team meeting	30 April 1996
10. Refinement of test protocols	May 1996
11. Training of laboratory personnel	May 1996
12. Distribution of first set of 10 coded chemicals	June 1996
13. Definition of data analysis procedures	August 1996
14. Preliminary data analysis (first 10 chemicals)	September 1996
15. 3rd Management Team meeting	16 September 1996
16. Distribution of remaining 50 chemicals	September 1996
17. Submission of results for second set of 50 chemicals	March 1997
18. Preliminary data analysis for all 60 test chemicals	June 1997
19. 4th Management Team meeting	18-20 June 1997
20. Statement on study outcome issued by Management Team	18 July 1997
21. 5th Management Team meeting	25 September 1997
22. Submission of reports for publication	October 1997
23. Reports accepted for publication in *Toxicology in Vitro*	April 1998
24. Endorsement of EPISKIN and TER by ESAC	April 1998
25. Draft EU Test Method submitted to EU National Coordinators	September 1998
26. Draft OECD Test Guideline submitted to OECD	September 1998
27. Acceptance of Test Method on skin corrosion by EU regulators	4 February 2000
28. Test Method on skin corrosion incorporated into Annex V of *Directive 67/548/EEC* by means of Commission *Directive 2000/33/EC*	25 April 2000
29. OECD Extended Expert Consultation on the draft OECD Test Guideline	1-2 November 2001
30. OECD circulates two new draft OECD Test Guidelines: TG 430 (TER) and TG 431 (human skin model) for comment	27 March 2002

3. THE SCIENTIFIC BASIS OF CHEMICAL RISK ASSESSMENT

3.1 The traditional risk assessment paradigm

The risk assessment of chemicals has traditionally been based on four elements:

a) hazard identification: the identification of the inherent capacity of a chemical to cause one or more adverse effects, without regard to the likelihood or severity of such effects;

b) hazard characterisation: the (semi-)quantitative evaluation of the nature of adverse effects following the exposure to a chemical, including the assessment of toxic potency (the relative toxicity of a chemical) and, where possible, a dose-response assessment;

c) exposure assessment: the (semi-)quantitative evaluation of the likely exposure of man and/or the environment to a chemical; and

d) risk characterisation: the (semi-)quantitative estimation of the probability that an adverse effect will occur, and of its severity and duration in a given population under defined exposure conditions, based on elements a, b and c.

3.2 Shortcomings of traditional risk assessment approaches

The different elements of the risk assessment paradigm encompass a variety of experimental activities. Hazard identification and characterisation (sometimes referred to collectively as hazard assessment) often rely on the use of animal experiments, whereas exposure assessment is generally the result of chemical analyses, but might also depend on biomonitoring in animals or humans and on computer-based estimations of exposure levels.

Animal-based studies generally lead to observations of the clinical, histopathological and/or functional changes in the animals caused by a given dose of the chemical under study. A common procedure is to determine the no observed adverse effect level (NOAEL) or the lowest observed effect level (LOEL). The NOAEL is the highest dose that produces no adverse effect in the most sensitive animal species, whereas the LOEL is the lowest dose that causes an adverse effect. An "adverse effect" can be defined as a change in the morphology, physiology, growth, development or lifespan of an organism that results in an impairment of its functional capacity or ability to compensate for additional stress, or in an increased susceptibility to the harmful effects of other environmental influences (IPCS, 1978). "Relevance" depends on the risk evaluation to be made: if the aim is to evaluate the risk of an adverse effect in humans, then the effect should be relevant to humans. Thus, an important question is whether the outcome of a particular animal experiment is relevant for the exposure of humans under practical circumstances.

In addition to ethical objections against, and legal restrictions on, the use of animals for toxicity testing, there are also scientific reasons for reducing the current reliance on animal data. The specific shortcomings of different animal tests have been discussed elsewhere (BUAV, 2001). More generally, there are difficulties associated with the need to extrapolate: a) across species, from relatively small, but

homogeneous, groups of laboratory animals to the very large and heterogeneous general human population; and b) from the high doses used to elicit effects in experimental animals to low doses, which are more consistent with human exposure levels. Thus, the use of animal data to predict the biological activities of compounds in humans is always prone to some degree of uncertainty. When assessing risk, toxicologists have traditionally attempted to overcome these uncertainties by introducing safety factors. For example, NOAELs determined from animal experiments are often divided by safety (uncertainty) factors to account for interspecies and/or inter-individual differences, when establishing safety standards for human exposure, such as an Acceptable Daily Intake (ADI; Kroes & Feron, 1990). Since it is considered acceptable to apply uncertainty factors when using animal data for risk assessment, it should also be acceptable to apply the concept of uncertainty when using *in vitro* data; for example, when extrapolating from cell culture systems to whole animals.

3.3 Advances in the scientific and technological basis for risk assessment

Over the last two decades, increasing emphasis has been placed on the development of non-animal test systems that are based on a fundamental biological understanding of the critical steps or events that link exposure to the expression of an adverse effect. The term "mode of action" is used to refer to one or more critical steps in the sequence of events that leads to a toxicological response, whereas the term "mechanism of action" refers to the complete cascade of events (EPA, 1999). In general, modes of action are better undersood than mechanisms of action.

Although the need for mechanistic tests is now widely recognised, it is not always clear what is meant by this term. For example, it could describe tests that involve biological systems with a mechanistic basis that is understood, or tests that are able to identify effects that are mechanistically related to the *in vivo* effects to be predicted. Frazier defined a mechanism as an explanation of an observed phenomenon that explains the processes underlying the phenomenon in terms of events at lower levels of organization (Frazier, 1994). Thus, a mechanistic test is based on a system at an acceptable level of organisation and a relevant endpoint based on a sufficient understanding of the cellular and/or molecular basis of the effect under consideration. An example would be a test based on interaction with a specific receptor, which is known to be a critical step in the development of a toxicological effect.

Scientific and technological advances can be recognised in a few key areas. Developments in molecular biology and cell culture techniques, in physiological measurements, and in genomics and proteomics are contributing significantly to our understanding of the basic biochemical and physiological processes underlying toxicological responses, and therefore to the scientific and technological basis for toxicity testing. In addition, developments in the use of computer-based modelling techniques ("*in silico* toxicology") are providing increasing more powerful ways of using and integrating non-animal data for predictive purposes.

Novel *in vitro* methods, such as those based on genetically-engineered cell lines, are being developed for the assessment of mechanistically-relevant endpoints. For example, the use of genetically-engineered mammalian cell lines with stable expression of specific human cytochrome P450 isoforms is permitting the study of

metabolic activation, inhibition and the effects human polymorphism (see Section 7.4.). Such cell lines have also been used for the assessment of specific target organ and target system toxicity. Examples include novel sytems for evaluating neurotoxicological hazard (Stingele *et al.*, 1999; Section 8.9) and embryotoxic hazard (Bremer *et al.*, 2001; Section 10.3).

Genomics and proteomics are areas of rapidly expanding knowledge. The application of genomics to toxicology is based on the assumption that most, if not all, toxic substances alter gene expression levels, so it should be possible to detect the potential toxicity of a chemical by screening for its ability to alter the expression of a diagnostic set of genes. In a study by Thomas *et al.* (2001), DNA microarray analysis was used to determine changes in the levels of gene transcripts expressed in the livers of mice exposed to five types of toxic chemicals: peroxisome proliferators, aryl hydrocarbon receptor agonists, noncoplanar polychlorinated biphenyls, inflammatory agents, and hypoxia-inducing agents. Statistical analysis of the gene patterns identified a set of 12 diagnostic transcripts, out of a total of 1200 genes, which enabled the toxic chemical category to be predicted with 100% predictive accuracy.

Another area of increasing knowledge concerns the relationship between the structures/properties of chemicals and their toxicities. For example, the association of specific molecular fragments with toxicological endpoints has led to the development of structure-activity relationships (SARs), whereas physicochemical properties, such as lipophilicity, hydrophilicity and molecular weight, have been used in the development of quantitative structure-activity relationships (QSARs). Knowledge of the physicochemical characteristics of a chemical are also important for an understanding of the biokinetic behaviour of the chemical. For example, quantitative structure-property relationships (QSPRs) have been used to predict biokinetic behaviour (DeJongh *et al.*, 1997, 1998; Soffers *et al.*, 2001).

A fourth area, commonly referred to as "*in silico* toxicology", has been the development of computer-based modelling techniques. Over the last fifteen years, the feasibility of computer-based approaches for predicting biokinetics has been greatly increased, due to the availability of computer techniques that allow for the simultaneous, numerical solution of differential equations. In addition, there has been substantial progress in the development of computational techniques that can be used for QSAR and toxicodynamic modelling.

As a result of these scientific and technological advances, new toxicity tests can now be developed on the basis of an improved mechanistic understanding of toxicological processes, and individual tests of different kinds can be combined in the form integrated testing strategies (Blaauboer *et al.*, 1999).

3.4 Approaches for reducing the amount of testing

When the REACH system is implemented, several approaches could be adopted for reducing the extent of animal and/or non-animal testing required to meet a given set of information requirements. In order to derogate from a standard set of assessment requirements, the use of read-across is already accepted by some regulatory authorities, and the concept of reverse risk assessment also deserves consideration.

3.4.1 Read-across

It can be assumed that chemicals with similar physicochemical property profiles will generally have similar toxicity profiles. Therefore, the clustering of chemicals with similar physicochemical property profiles into groups will permit "read-across" of toxicological properties within the groups, thus reducing the extent of testing required for chemicals within the group. Read-across is the process by which one or more toxicological properties of a given chemical are inferred by comparison of that chemical with chemicals of similar molecular structures and physicochemical properties, for which the toxicological properties of interest are known. The process involves an appraisal by a regulatory authority on the basis of information provided by industry.

A number of issues need to be considered when assessing the toxicological properties of a new substance by read-across: a) the similarity of the purity and impurity profiles of the new substance and the structural analogue needs to be assessed, since there should be no differences in the purity or impurities in a scale that would be likely to influence the overall toxicity; b) the physicochemical properties of the new substance should be compared with its analogue; in particular, the physical form, molecular weight, water solubility, partition coefficient and vapour pressure provide useful information as to similarity; c) the likely toxicokinetics of the substances, including the possibility of different metabolic pathways, should be considered; d) if read-across data on structural analogues have not been produced by using current Annex V test methods or current OECD test guidelines, caution should be exercised when extrapolating these data to a new chemical; and e) regulatory authorities are more likely to accept the read-across of positive findings (presence of toxicity), rather than negative findings (absence of toxicity).

As an illustration of the use of read-across in a regulatory submission for new chemicals, a request received by UK Health and Safety Executive involved a series of four structurally-similar substances differing in numbers of carbon atoms. The result was full base-set testing of the low molecular weight member of the series and limited testing on the highest molecular weight member. For the other group members, all the toxicological data used for base-set notification were read-across data (Rosalind Hanaway, personal communication). Further information on the use of read-across in a regulatory context is given by Hanway & Evans (2000).

Other hypothetical examples of read-across are provided by Barratt (2000). For example, in a homologous series of surfactants (alkyl betianes) with 8, 10, 12, 14 and 16 carbon alkyl chains, the systemic toxicities of the 10-carbon and 14-carbon chemicals could be predicted from tests carried out on the 8, 12 and 16-carbon members of the series. This would result in a 40% reduction in animal usage.

3.4.2 Reverse risk assessment

A "reverse risk assessment" enables the extent of testing needed for a chemical to be judged according to its prospective use and exposure conditions, instead of the performance of a standard list of hazard assessment tests followed by a risk assessment. In the latter situation, the margin of safety eventually obtained could be so high that it would be realised that much of the testing was unnecessary in the first

place. The margin of safety is the magnitude by which the NOAEL exceeds the known or estimated exposure level.

The first step is to perform an exposure assessment, based on experimental measurements and/or predictions. Once the likely exposure is known or predicted, a NOAEL (or LOEL) is derived by measuring organ-specific cytotoxicities. If the likely exposure is significantly less than the NOAEL (i.e. the chemical has a lower toxicity than would cause concern in the likely exposure) then no more testing is required. Since the NOAEL is obtained by using isolated cells, this implies that 100% of the chemical is absorbed. If the NOAEL is of the same order or slightly less than the exposure level, then the possibility that absorption is less than 100% should be considered in deciding whether further testing is necessary. Conversely, if the NOAEL is significantly less than the exposure level, it may be necessary to undertake additional testing.

A proposal for an exposure-driven risk assessment process has been developed by the European oil industry organisation for environment, health and safety, CONCAWE (unpublished document submitted to the Commission Working Group on Testing, Registration and Evaluation). According to this proposal, the extent of testing considered necessary is determined by using an objective measure of exposure, based on tonnage, use category and physicochemical properties.

3.5 The use of alternative methods in hazard and risk assessment

3.5.1 The use of physicochemical data by QSAR

Physicochemical data are useful at an early stage in the hazard/risk assessment process, since it may be possible to extrapolate them to toxicological effects by means of QSAR models (in which the descriptor variables are physicochemical properties). In the context of the future chemicals policy, QSARs based on the physicochemical properties required for the base-set notification would be particularly useful. These properties are summarised in Table 3.1.

3.5.2 The use of in vitro data

The use of *in vitro* tests in different components of the risk assessment process (hazard identification, hazard characterisation and risk assessment) was discussed by Balls & Fentem (1992). As described in Chapter 2.2, the data generated by *in vitro* tests can be extrapolated to toxicological effects by using prediction models (PMs). In principle, various PMs can be applied to the data generated by a particular *in vitro* test. Classification models can be used to identify chemicals that are potentially hazardous (hazard identification), and to predict their regulatory classifications (hazard classification). Examples include PMs for predicting skin corrosion potential (Fentem *et al*., 1998), such as the following PM based on pH measurements (OECD, 1998):

If the pH of a substance ≤ 2, or if pH ≥ 11.5, classify as a corrosive.

If different classification systems exist for a particular type of toxic hazard (e.g. the EU system and the Globally Harmonised System [GHS]), different PMs might need to be applied to the same *in vitro* data. For the purposes of risk management, the

classification of chemicals might be considered to be a sufficient basis for making certain decisions, such as on fast-track risk reduction measures.

In other cases, it might be necessary to perform a more-comprehensive hazard characterisation; for example, to obtain an estimate of toxic potency. In such cases, a different kind of PM is required, which could be a regression model. An example is a PM for estimating acute lethal doses (LD50 values) from cytotoxicity (IC50) data (Spielmann *et al.*, 1999).

Finally, if a risk assessment is required, *in vitro* dose-response data would need to be used in combination with exposure measurements or estimates, taking into account biokinetic considerations (see Chapter 7 for further details).

3.5.3 The integrated use of physicochemical data and in vitro data

Finally, it is recommended that the assessment of toxicological endpoints is based on the intelligent and combined use of physicochemical and *in vitro* data, in the form of tiered strategies. Examples include the tiered testing strategies for acute dermal and ocular toxicity, described in Chapter 5. The development and validation of such strategies should be a matter of high priority, since they combine the strengths of complementary approaches.

Table 3.1. Physicochemical properties required for the base-set notification.

No.	Physicochemical property
1	Melting point (melting range) *
2	Boiling point (boiling range) *
3	Relative density
4	Vapour pressure *
5	Surface tension
6	Water solubility
7	Fat solubility
8	Partition coefficient (logP)
9	Flash point *
10	Flammability
11	Explosive properties
12	Auto-flammability *
13	Oxidising properties
14	Abiotic degradation: hydrolysis as a function of pH (a test requirement for ecotoxicity)

* The physical state of the chemical may make the determination of some of these properties inappropriate.

4. ACUTE LETHAL TOXICITY

4.1 Introduction

Acute systemic toxicity testing is conducted to determine the hazard potential of a single exposure to a chemical or product (by the oral, dermal or respiratory routes). The assessment of lethality, which forms part of acute systemic toxicity testing, has traditionally relied on the *in vivo* measurement of the median lethal dose (LD50), i.e. the dose of a test chemical that kills 50% of the animals in the experimental group. During acute toxicity testing, non-lethal endpoints can also be assessed, to identify specific target organ and target system toxicities.

The LD50 value is used as a reference value to classify chemicals on the basis of acute toxic hazard, and to define appropriate doses for use in other *in vivo* tests. The classical LD50 test attracted much criticism for both scientific and animal welfare reasons (Zbinden & Flury-Roversi, 1981), and was modified to become an approximate LD50 test or a limit test. More recently, reduction and refinement alternatives have been accepted at the OECD level (OECD, 1992; 1996; 1998).

The main question being addressed in the development and validation of alternatives of acute lethality testing is whether QSAR and/or *in vitro* methods can provide a reliable basis for acute hazard classification and/or a valid estimate of appropriate starting doses for non-lethal *in vivo* studies.

4.2 *In vitro* methods for acute lethal toxicity

Acute systemic toxicity *in vivo* can result from toxicity at the cellular level, i.e. cytotoxicity, which in turn results from interference with structures and/or properties essential for cell survival, proliferation and/or function. These effects can involve, for example, the integrity of membranes and the cytoskeleton, metabolism, the synthesis and degradation or release of cellular constituents or products, ion regulation, and cell division. It is useful to distinguish between three types of cytotoxicity (Seibert *et al.*, 1996). *Basal (or general) cytotoxicity* involves one or more of the above-mentioned structures or processes, when all of the cell types studied show similar sensitivities. *Selective (or cell-specific) cytotoxicity* occurs where some types of differentiated cells are more sensitive to the effects of a particular toxicant than others; for example, as a result of biotransformation, binding to specific receptors, or uptake by specific mechanisms. *Cell-specific function toxicity* occurs when the toxicant affects structures or processes which may not be critical for the affected cells themselves, but which are critical for the organism as a whole. For example, such toxicity can involve effects on cell-cell communication, via the synthesis, release, binding and degradation of cytokines, hormones and transmitters, or on specific transport processes.

A large number of *in vitro* cytotoxicity tests have been developed, employing a variety of cell lines and endpoint measurements (Table 4.1).

The Fund for the Replacement of Animals in Medical Experiments (FRAME) cytotoxicity (BCL-D1 Kenacid Blue) assay went through a blind trial involving four

laboratories and 50 chemicals (Knox *et al*., 1986). Subsequently, the same chemicals were put through the 3T3 Kenacid Blue assay (Clothier *et al.*, 1998). These studies demonstrated the reproducibility of the cytotoxicity assays concerned.

The focus of this chapter is on the assessment of basal cytotoxicity and its association with acute lethal toxicity. The rationale for relating basal cytotoxicity to acute lethal toxicity is that the latter is often due to critical organ failure, which results from loss of cell viability and function. No basal cytotoxicity tests have been through a formal validation study for the specific purpose of replacing the rodent LD50 test, although a number of *in vitro-in vivo* correlation studies have indicated that IC50 values can reliably be used to predict LD50 values with a reasonable degree of precision (Table 4.2). An example of such a study on 59 chemicals is provided by Clothier *et al.* (1987).

4.3 *In vitro* cytotoxicity databases

A huge amount of *in vitro* cytotoxicity data exist, some of which have been incorporated into the following databases, which are therefore useful resources for test development and validation activities:

1. The Halle and Gores Registry of Cytotoxicity, which contains 1,912 individual IC50 values for 347 chemicals. The IC50 values were obtained from multiple reports in the literature, were averaged for each chemical, and then paired with acute oral LD50 values for the rat, and with intra-peritoneal (i.p) LD50 values for the mouse, obtained from the National Institute for Occupational Safety and Health (NIOSH) Registry of Toxic Effects of Chemical Substances (RTECS). The details of this registry were originally published in German (Halle & Goeres, 1998), but have been translated into English under the auspices of ECVAM and ZEBET, and will shortly be published in *ATLA.*

2. The FRAME database, which contains IC20, IC50 and IC80 values for over 250 chemicals tested on mouse 3T3 cells with the Kenacid Blue assay for total protein content. Fifty-nine IC50 values have been compared with mouse i.p. and rat oral LD50 data.

3. The Multi-centre Evaluation of *In Vitro* Cytotoxicity (MEIC) database, which contains data for 50 chemicals and 69 different assay methods (Ekwall *et al.*, 1998). In the MEIC study, the data were compared with human lethal blood concentrations (Clemedson *et al.*, 2000; Scheers *et al.*, 2001).

4. The MEMO (the MEIC monographs on time-related human lethal blood concentrations) database, which supplements the MEIC database with data for an additional 25 additional chemicals.

4.4 The MEIC and EDIT projects

The results of the MEIC study demonstrated that human basal cytotoxicity tests are relevant for predicting the human acute toxicity of chemicals. In addition, the results show that other important toxic mechanisms exist, which only might be measured by

supplementary *in vitro* toxicity tests, and that modelling of human toxicity is improved by the additional use of toxicokinetic data.

The EDIT (Evaluation-guided Development of New *In Vitro* Tests) project was initiated as a follow-up to the MEIC study in 1998 (Ekwall *et al.*, 1999). The EDIT project was established as a six-year effort by international cytotoxicology laboratories, with the aim of developing new *in vitro* tests for toxicity and toxicokinetics, for incorporation into optimised test batteries for acute and chronic systemic toxicity.

4.5 The international workshop on *in vitro* methods for acute systemic toxicity

The status of several major international *in vitro* initiatives directed toward reducing the use of laboratory animals for acute toxicity testing was reviewed in October, 2000, at the International Workshop on *In Vitro* Methods for Assessing Acute Systemic Toxicity (NIH, 2001a). The workshop participants concluded that none of the proposed *in vitro* models had been validated for reliability and relevance in any formal studies. The participants made recommendations for future research and development, and for validation efforts in the field of *in vitro* systemic toxicity testing.

One of the workshop recommendations for reducing and refining the use of animals for lethality assays in the near-term was the publication of guidance for using *in vitro* cytotoxicity assays to estimate starting doses for acute oral lethality assays (NIH, 2001b). The use of such assays could reduce animal testing by obviating the need to perform dose range-finding experiments in animals.

As a follow-up to the international workshop, there needs to be a validation study the use of the murine 3T3 cells and human keratinocytes for predicting acute systemic toxicity. The study could initially lead to the definition of a standardised protocol (including positive and negative controls) and could define limiting parameters for establishing the reliability of the method (Spielmann *et al*., 1991; Triglia *et al*., 1989). The validation study could be used to evaluate the capacity of the basal cytotoxicity assays to predict: a) rat oral and/or mouse i.p. LD50 values; and b) human lethal blood concentrations. To meet these goals, a collaborative validation study is currently being planned by ICCVAM and ECVAM.

4.6 Structure-activity relationships for acute lethal toxicity

QSARs for acute lethal toxicity to rodents have been reviewed by Cronin & Dearden (1995). More recently, Johnson & Jurs (1997) used regression analysis and neural networks to predict the rodent toxicity of a diverse set of 115 anilines. In a study by Wang *et al.* (1999), the LD_{50} values of over 46,000 single organic substances taken from RTECS database were analysed, on the basis of which a number of structural alerts (i.e. fragments associated with toxicity) were identified. In a study by Calleja *et al.* (1994), QSARs for acute lethal toxicity to humans were investigated by using data for 38 structurally-diverse chemicals, selected from the 50 priority chemicals identified in the MEIC programme (see below). The most predictive physicochemical properties were the octanol-water partition coefficient (logP) and heat of formation, along with additional properties which describe molecular size and electronic

properties. The results of such studies are promising, but more work will be needed to assess the extent to which QSAR models can be used to replace and/or reduce *in vivo* testing. One possibility is that QSAR models could be used to prioritise the *in vitro* testing of a number of chemicals, or to decide on the need to perform *in vitro* testing for a particular chemical.

4.7 A tiered testing strategy for acute lethal toxicity

The outline of a testing strategy for the classification and labelling of chemicals according to their acute systemic toxicity was proposed in ECVAM Workshop Report 16 (Seibert *et al.*, 1996), and is developed further in Figure 4.1. The first step is based on the use of QSARs for predicting acute toxicity. If the result of this step indicated that the compound should be classified as "very toxic", no testing would need to be performed. The second step is based on a basal cytotoxicity test, such as the 3T3 neutral red uptake (NRU) test. If the result of the second step indicated that the compound should be classified as "very toxic", no further testing would need to be done. If not, the third step, based on the assessment of biotransformation, would have to be performed. Again, if the result were positive, testing would be stopped at this stage. If not, the third step, based on the use of cell-specific toxicity tests, would have to be performed. Finally, the chemical would be classified as "very toxic", "toxic", "harmful" or "no label", according to the lowest EC50 value determined at any of the three testing levels. If all of the results indicated that the chemicals should be assigned to the lowest toxicity class (i.e. "no label"), a limited *in vivo* study might need to be carried out, to confirm the absence of significant acute lethal potential.

4.8 Acute lethal toxicity: summary, conclusions and recommendations

Standardised basal cytotoxicity tests, such as the 3T3 NRU assay, are already widely used for non-regulatory purposes, and could be used immediately for priority setting among chemicals, and for establishing the starting dose for *in vivo* acute toxicity testing.

Future activities in this area should aim to: a) reduce and replace the use of the rodent test for determining LD50 values, through the development and validation of QSARs and cytotoxicity tests; b) ensure the widest possible use of QSARs and cytotoxicity tests as a means of estimating LD50 values; and c) promote the development and validation of a tiered testing strategy for acute systemic toxicity.

The prospects outlined below focus on the use of QSARs and basal cytotoxicity tests for predicting acute lethal toxicity. The uses of alternative methods for predicting metabolism (Step 3 in Figure 4.1) and organ-specific toxicities (Step 4 in Figure 4.1) are addressed in chapters 7 and 8, respectively.

4.8.1 Short-term prospects

1. Development of a comprehensive database containing *in vitro* results (e.g. IC50 values) and corresponding *in vivo* LD50 values (rat oral or mouse i.p, and/or human lethal blood levels), to allow *in vitro-in vivo* comparisons to be made. ECVAM has initiated the development of such a database.

51

2. Validation of basal cytotoxicity assays for predicting: a) rat oral and/or mouse i.p. LD50 values; and b) human lethal blood concentrations. Such a study has been initiated under the auspices of ICCVAM and ECVAM.
3. Development of QSAR models for predicting *in vitro* cytotoxicity.

4.8.2 Medium-term prospects

1. Validation of QSAR models for predicting *in vitro* cytotoxicity.
2. Replacement of the need for *in vivo* testing for classifying chemicals on the basis of acute toxic potential, as well as for prioritising and dose-setting for other kinds of studies.

Figure 4.1. Tiered testing strategy for acute lethal toxicity.

Step 1:   Use a validated QSAR model   ⇒   classify as "very toxic"
                                                              and stop testing

          ⇓
          not very toxic
          ⇓

Step 2:   Perform a validated basal cytotoxicity test
          and apply a validated prediction model   ⇒   classify as "very toxic"
                                                              and stop testing

          ⇓
          not very toxic
          ⇓

Step 3a: Use a validated computer model
          for metabolism   ⇒   classify as "very toxic"
                                                              and stop testing

          ⇓
          not very toxic
          ⇓

Step 3b: Perform validated biotransformation tests
          for metabolism   ⇒   classify as "very toxic"
                                                              and stop testing

          ⇓
          not very toxic
          ⇓

Step 4:   Perform validated cell-specific toxicity   ⇒   classify as "very toxic",
          tests and apply validated prediction models   "toxic" or "harmful", and
                                                              stop testing

          ⇓
          not toxic ("no label")
          ⇓

Step 5:   Conduct a limited i*n vivo* acute toxicity
          study   ⇒   classify as "no label" or
                                                              assign to appropriate class

Footnotes

To interpret the results of step 3 (biotransformation assessment) in terms of an LD50 estimate, it would be necessary to apply a suitable prediction model or biokinetic model.

Animal testing (step 5) should not be conducted as a matter of routine, but only in situations where a staisfactory classification cannot be derived by the application of steps 1-4.

Table 4.1. An overview of *in vitro* assays for cytotoxicity / acute lethal toxicity.

Endpoint	Measurement	Cell Line(s)	Study
Cell growth/viability	ATP content or leakage	ELD cells (mouse); erythrocytes (mouse); LS-L929 cells (mouse); hepatocytes (rat); spermatozoa (bovine); HL-60 cells (human)	Clemedson *et al.*, 2000
	Cell morphology	C9 cells (rat); hepatocytes (rat); L2 cells (rat); MDBK cells (bovine); Chang liver cells (human); HeLa cells (human); McCoy cells (human); WI-1003/Hep-G2 cells (human)	Clemedson *et al.*, 2000
	Chromium release	LS-L929 cells (mouse)	Clemedson *et al.*, 2000
	Creatine kinase activity	Muscle cells (rat)	Clemedson *et al.*, 2000
	Haemolysis	Erythrocytes (human)	Clemedson *et al.*, 2000
	Killing index (sic)	SQ-5 cells (human)	Clemedson *et al.*, 2000
	LDH release	3T3 cells (mouse); hepatocytes (rat, human); Hep-2 cells (human); Hep-G2 cells (human); lymphocytes (human); SQ-5 cells (human); corneal cells (human)	Clemedson *et al.*, 2000
	Neutral red uptake	3T3 cells (mouse); L929 cells (mouse); V79 cells; NB41-A3 cells (mouse); BHK cells (hamster); hepatocytes (rat, human); HeLa cells (human) Hep-2 cells (human); keratinocytes (human)	Start *et al.,* 1986 Riddell *et al.*, 1986a,b Clothier *et al.*, 1999 Spielmann *et al.*, 1999 Clemedson *et al.*, 2000
	Neutral red release	Rabbit corneal fibroblasts or mouse embryonic fibroblasts or normal human epidermal keratinocytes	Reader *et al.*, 1989
	Plating efficiency	HeLa cells (human	Clemedson *et al.*, 2000
	Viable cell count	LS-L929 cells (mouse); polymorphonuclear leukocytes (human)	Clemedson *et al.*, 2000
	Cell cycle distribution	Daudi cells (human); RERF-LC-AI cells (human)	Clemedson *et al.*, 2000
	Glucose consumption	Muscle cells (rat)	Clemedson *et al.*, 2000
	Macromolecule content	HTC cells (rat); Hep-G2 cells (human)	Clemedson *et al.*, 2000
	MTT metabolism	3T3 cells (mouse); L929 cells (mouse); NG108-15 cells (mouse, rat); V79 cells (hamster); hepatocytes (rat, human); Detroit 155, DET dermal fibroblast	Clemedson *et al.*, 2000

		(human); FaO cells (human); HFL1 cells (human); 3D Skin[2], Dermal Model ZK1100 keratinocytes (human); lymphocytes (human); RERF-LC-AI cells (human); WS1 cells (human)	
Cell growth	pH change	L2 cells (rat); Chang liver cells (human); HeLa cells (human); WI-1003/Hep-G2 cells (human)	
	Total protein content	3T3 or 3T3-L1 cells (mouse); Hepa-1c1c7 (mouse); L929 cells (mouse); V79 cells (hamster); hepatocytes (rat); PC12h cells (rat); LLC-PK1 cells (pig); HeLa cells (human); MRC-5 cells (human); NB-1 cells (human); Chinese hamster V79 cells	Hulme *et al.,* 1987 Clothier *et al.*, 1988 Fry *et al.*, 1990 Spielmann *et al.,* 1999 Clemedson *et al.*, 2000
	Tritiated-proline uptake	L2 cells (rat)	Clemedson *et al.*, 2000
	Tritiated-thymidine incorporation	Peripheral lymphocytes (human)	Spielmann *et al.*, 1999 Clemedson *et al.*, 2000
Specialised function effects	Cell resting membrane potential	NG108-15 (mouse, rat)	Clemedson *et al.*, 2000
	Chemotactic peptide-stimulated chemotaxis / locomotion	Polymorphonuclear leukocytes (human)	Clemedson *et al.*, 2000
	Ethoxycoumarin deethylase (EOD) activity	Hepatocytes (rat)	Clemedson *et al.*, 2000
	Inhibition of NK cell-mediated cytotoxicity	Natural killer cells (human)	Clemedson *et al.*, 2000
	Intracellular glycogen content	Hepatocytes (rat)	Clemedson *et al.*, 2000
	Motility or velocity	Spermatozoa (bovine)	Clemedson *et al.*, 2000
	Spontaneous contractility	Muscle cells (rat)	Clemedson *et al.*, 2000

Footnote

ATP=adenosine triphosphate; LDH=lactate dehydrogenase; MTT=3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide; NK=natural killer.

Table 4.2   Acute systemic toxicity studies and/or approaches

Most relevant *in vitro/in vivo* correlation studies and/or approaches	References
Studies conducted by FRAME and partners	Knox *et al*., 1986 Riddell *et al.*, 1986a,b Clothier *et al*, 1987 Hulme *et al.*, 1987 Fry *et al.*, 1990 Balls *et al.*, 1992 Fentem *et al.*, 1993 Clothier *et al.*, 1998
The MEIC (Multicentre Evaluation of In Vitro Cytotoxicity) study	Clemedson & Ekwall, 1999 Ekwall *et al.*, 2000
The ECITTS (ERGATT/CFN Integrated Toxicity Testing Strategy) study	De Jongh *et al.*, 1999
Japanese Society of Alternatives to Animal Experiments (JSSAE) activities	Ohno *et al.*, 1998a; 1998b; 1998c
Testing framework proposed under the auspices of SGOMSEC	Curren *et al.*, 1998
General form of a testing strategy proposed during an ECVAM workshop	Seibert *et al.*, 1996
The ZEBET approach for predicting starting doses for acute *in vivo* testing	Halle & Goeres, 1988 Spielmann *et al.*, 1996 Spielmann *et al.*, 1999 Halle *et al.*, 2000

5. LOCAL TOXICITY: ACUTE DERMAL AND OCULAR EFFECTS

5.1 Introduction

Acute local toxicity refers to the local toxic effects that may result from a single exposure to a chemical or product, via the oral, dermal, ocular or inhalation route. The exposure can be incidental, accidental or deliberate (for example, in the case of cosmetics and certain medicines). The dermal route of exposure can also be a significant route for entry into the systemic circulation, as discussed in the section on percutaneous absortpion (Section 7.2). The main types of toxic effects are dermal and ocular corrosion and irritation, and these are generally assessed in a sequential manner in the context of tiered assessment strategies, such as those adopted by the OECD (2001). In such strategies, dermal effects are generally assessed before ocular effects. Tiered approaches for assessing corrosion and irritation are illustrated in Figures 5.1 (dermal effects) and 5.2 (ocular effects).

5.2 Current status of alternative methods for skin corrosion

5.2.1 Structure-activity relationships for skin corrosion

Various SARs for skin corrosion have been reported by Barratt and colleagues (Barratt, 1996a; Barratt, 1996b; Whittle *et al.*, 1996; Barratt *et al.*, 1998). On the whole, the SARs presented in these studies take the form of principal component (PC) plots, which are based on physicochemical properties and show a separation between corrosive (C) and non-corrosive (NC) chemicals. Explicit classification models were not presented. Rather than modelling a heterogeneous group of chemicals, separate analyses were performed for acids, bases, electrophiles and neutral organics (defined as uncharged molecules which lack the potential to react covalently and which do not ionise under biological conditions [Martin Barratt, personal communication]). The most recent presentation of this approach is given in Barratt *et al.* (1998). In addition to PC analyses, discriminant analysis and neural network analysis were also applied to a group of neutral and electrophilic chemicals (Barratt, 1996a), and to the acids, bases and phenols (Barratt, 1996b). Finally, in another study (Barratt *et al.*, 1996), PC plots for acids were based not only on physicochemical propertes, but also on *in vitro* cytotoxicity measurements in mouse 3T3 cells. More recently, it was shown that a heterogeneous set of organic chemicals could be predicted as C or NC on the basis of melting point (Mpt) and molecular weight (MW) (Worth, 2000), according to the following PM:

If Mpt ≥ 37 °C and MW ≤ 123 g/mol, predict as C; otherwise predict as NC.

Similar rules have been developed by Gerner and colleagues, who have incorporated a system of decision rules into an expert system used by the German BgVV (Gerner *et al.*, 2000a, 2000b; Zinke *et al.*, 2000). An example is the following PM:

If MW > 1200 g/mol, then the substance has no local toxic effects.

5.2.2 In vitro methods for skin corrosion

The current status of alternative methods is summarised in Table 5.1. ECVAM-funded validation studies on *in vitro* tests for skin corrosion have been conducted (Fentem *et al.,* 1998; Liebsch *et al.,* 2000), and the scientific validities of four *in vitro* tests have been endorsed by the ESAC: the rat skin transcutaneous electrical resistance (TER) assay (ECVAM, 1998), two tests based on the use of commercial reconstituted skin equivalents, EPISKIN™ (ECVAM, 1998) and EpiDerm™ (ECVAM, 2000), and Corrositex™ (ECVAM, 2001).

In the EU, a new Test Method on Skin Corrosion, incorporating the rat skin TER and human skin model assays, has been included in Annex V of the Dangerous Substances Directive (*Directive 67/548/EEC*; EC, 2000), thereby making the use of *in vitro* alternatives for testing the skin corrosion potential of chemicals mandatory in the EU.

A draft Test Guideline (TG) on *in vitro* tests for skin corrosion was submitted to the OECD in late 1998, for consideration by the OECD Member Countries. Following a number of commenting rounds, an expert meeting, held on 1-2 November 2001 in Berlin, agreed that the draft TG on *in vitro* skin corrosion should be divided into two separate TGs: a draft proposal for a new TG 430 on the TER test (not restricted to the rat skin TER test) and a draft proposal for a new TG 431 on the human skin model test. The new TG 430 and TG 431 are now being considered by OECD Member Countries.

5.2.3 Tiered testing strategies for skin corrosion

In 1996, the OECD proposed a tiered (stepwise) approach to hazard identification, which underwent revisions in 1998 and in 2001 (OECD, 2001). In the EU, a tiered testing strategy for skin corrosion/irritation is being proposed for incorporation into Annex V of *Directive 67/548/EEC*. This could be achieved during 2002 by means of the 29th Adaptation to Technical Progress (ATP) of the directive (Juan Riego-Sintes, personal communication).

An evaluation of a two-step strategy, based on the sequential use of pH measurements and *in vitro* data, indicated that the use of pH data in addition to TER or EPISKIN data, improves the ability to predict corrosion potential (Worth & Cronin, 2001a). An evaluation of a three-step strategy, based on the sequential use of QSARs, pH measurements and *in vitro* data, indicated that tiered approaches provide an effective means of classifying chemicals, while at the same time reducing and refining the use of animals (Worth *et al.,* 1998). A study carried out by ECVAM confirmed the usefulness of pH as a predictor of skin corrosion potential, and provided a new PM for identifying chemicals that are corrosive by a pH-dependent mechanism (Worth & Cronin, 2001b)

5.3 Skin corrosion: summary, conclusions and recommendations

Alternative methods for skin corrosion have been validated and accepted for regulatory use in the EU, so animal testing should not be performed for this endpoint. The hazard identification (classification and labelling) of skin corrosives should be based on the use of a pH test, where appropriate, and an *in vitro* test (rat skin TER assay, human skin model assay or, for qualifying test chemicals, CORROSITEX®).

For risk assessment (dose-response investigations, coupled with assessments of skin irritation potential at doses negative in skin corrosion tests), the rat skin TER or a human skin model assay are recommended for use.

5.3.1 Short-term prospects

1. The achievement of acceptance by OECD Member Countries of the new TGs on *in vitro* tests for skin corrosion, although this is not a pre-requisite for the use of these tests in the context of chemicals testing in the EU.
2. The validation of QSARs and/or expert system rulebases for skin corrosion.

5.4 Current status of alternative methods for skin irritation

5.4.1 Structure-activity relationships for skin irritation

Relatively few QSAR studies for skin irritation have been reported in the literature. Barratt (1996a) reported a QSAR for predicting the primary irritation index (PII) of organic chemicals, but this had little predictive value ($r^2 = 0.42$). In the same study, discriminant analysis was shown to discriminate between irritant and non-irritant chemicals, as defined by EU classification criteria, with an accuracy of 67%.

Hayashi *et al.* (1999) reported two QSARs for predicting the molar-weighted PII of phenols. One model, based on absolute hardness, was proposed for chemicals with negative lowest unoccupied molecular orbital (LUMO) energies, whereas the other model, based on the logarithm of the octanol-water partition coefficient (logP), was proposed for chemicals with positive LUMO energies. These models had correlation coefficients of 0.72 and 0.82, respectively (i.e. r^2 values of 0.52 and 0.67).

Smith *et al.* (2000) analysed a data set of 42 esters, for which human skin irritation data were available, and for which 19 physicochemical properties had been calculated. Best subsets regression was used to select variables for subsequent inclusion in discriminant models. The best variables were water solubility (lower for irritants than non-irritants), a dispersion parameter (higher for irritants), a hydrogen-bonding parameter (higher for irritants), the sum of partial positive charges (lower for irritants), and density (lower for irritants). A discriminant model based on all five parameters had a sensitivity of 85% and a specificity of 92%.

5.4.2 In vitro methods for skin irritation

At present, there are no validated *in vitro* tests for skin irritation. Acting on a recommendation by the ECVAM Skin Irritation Task Force (Botham *et al*., 1998), ECVAM has supported prevalidation studies on five *in vitro* tests for acute skin irritation: EpiDerm™, EPISKIN™, PREDISKIN™, the pig ear test, and the mouse skin integrity function test (SIFT) (Fentem *et al.,* 2001; Heylings *et al.,* 2001). The outcome of these studies was that none of the tests was ready for progression to formal validation, but that appropriate modifications to certain test protocols might enable them to meet the criteria for early inclusion in a future validation study.

Various follow-up activities to the prevalidation study are currently in progress, with the objective of having test protocols suitable for inclusion in a formal validation study within 2002. An extended ECVAM Skin Irritation Task Force meeting was held in May 2001, to agree and prioritise the activities required prior to setting up a

validation study (Zuang *et al.,* 2002). On the basis of additional work on the EPISKIN™, EpiDerm™ and SIFT test protocols and/or prediction models, conducted subsequent to the prevalidation studies, it is hoped that a common EpiDerm/EPISKIN protocol could meet the performance criteria defined for progression to a validation study.

Several testing strategies have been proposed/evaluated (Botham *et al.,* 1998; OECD, 2001; Robinson *et al.,* 2000), some of which involve human volunteer studies (human 4-hour patch test; Basketter *et al.,* 1997). While it may be appropriate to conduct such studies on a case-by-case basis for risk assessment purposes, human testing has no role in the early screening and hazard identification of chemicals.

5.4.3 In vitro methods for phototoxicity

Phototoxicity is a broad term covering a number of endpoints, including acute photoirritation, photogenotoxicity and photocarcinogencity. Although not currently part of the testing requirements for chemicals, some phototoxicity endpoints are required for the assessment of cosmetic and pharmaceutical products.

Acute photoirritation refers to the toxic response that may be induced in the skin upon exposure to light, following earlier exposure to certain chemicals either by the topical or the systemic route (Spielmann *et al.*, 1994, 2000). The 3T3 NRU assay for acute photoirritation has been successfully validated (Spielmann *et al.*, 1998a, 1998b) and endorsed by the ESAC (ECVAM, 1998a, 1998b). A test guideline based upon the validated protocol has been accepted as an Annex V method in relation to *Directive 67/548/EEC*, and is currently being considered in the OECD Test Guidelines Programme for acceptance as an OECD Test Guideline.

Several approaches have been developed for photomutagenicity testing. For example, Dean *et al.* (1991) proposed the combined use of a bacterial mutation test in *E.coli* and a mammalian cytogenetics assay in Chinese Hamster Ovary (CHO) cells. Photomutagenicity tests based on bacterial, yeast and mammalian cells have also been proposed by Chetelat *et al.* (1993a, 1993b).

5.5 Skin irritation: summary, conclusions and recommendations

In vitro methods for skin irritation testing could be used immediately for priority setting. The human skin model assays (e.g. EpiDerm™ and EPISKIN™) and the mouse SIFT appear to be the most promising (Zuang *et al.*, 2002).

For risk assessment purposes, there is a need to identify and evaluate the usefulness of new, mechanistically-based endpoints that are more predictive of skin irritation than are simple cytotoxicity determinations. The existing *in vitro* models also need to be improved, so that they are better representative of the skin *in vivo*.

5.5.2 Short-term prospects

1. The validation of modified test protocols for the human skin models and SIFT, to determine whether of any of these existing methods can adequately distinguish acute skin irritants from non-irritants. The key activities required prior to organising a validation study, which should be coordinated by ECVAM, are:

a) identification of test chemicals (irritants, non-irritants) for use in (pre)validation studies from the New Chemicals Database (ECB, JRC Ispra); and

b) assessment under blind conditions of a common protocol for the EPISKIN™ and EpiDerm™ tests, and of a modifed protocol for the SIFT test, to check whether these are ready for formal validation. It is envisaged that this will be a small study in a single laboratory per test, with approximately 20 chemicals not previously used in this work.

2. The further development of (Q)SARs and/or expert system rulebases for skin irritation.

3. The acceptance during 2002 by the OECD Council of a revised TG 404 (acute dermal irritation/corrosion), which includes as a supplement the OECD tiered testing strategy for skin irritation and corrosion.

5.5.3 Medium-term prospects

The validation of QSARs and/or expert system rulebases for skin irritation.

5.5.4 Recommendations for research and development

More resources should be provided for research aimed at the identification of new markers for skin irritation, and on-going activities should be coordinated, with a view to having identified promising toxicological endpoints and developed new toxicity tests for validation by 2006. This research should be undertaken in parallel with the validation of existing test protocols for hazard identification. One approach to this research is through the application of genomics and proteomics. For example, COLIPA has started a three-year programme on specific aspects of proteomics and genomics.

5.6 Current status of alternative methods for eye irritation and corrosion

5.6.1 In vitro methods for eye corrosion

The tests for skin corrosion described above are assumed to identify chemicals that would also be corrosive to the eye. This should be clearly spelled out in Annex V of *Directive 67/548/EEC*, in the description of the methods for skin corrosivity testing (EC, 2000).

5.6.2 In vitro methods for eye irritation

Six major validation or evaluation studies took place from 1991 to 1997:

1) the EC/HO study (Balls *et al.*, 1995);
2) the COLIPA study (Brantom *et al.*, 1997);
3) the BGA/BMBF study (Spielmann *et al.*, 1991, 1993, 1996);
4) the CTFA study (Gettings *et al.*, 1991, 1992, 1994, 1996);
5) the IRAG study (Bradlaw *et al.*, 1997);
6) the MHW/JCIA (Ohno *et al.*, 1994) study, in which in total about 44 alternative methods (not counting variations on the various methods) have been evaluated.

The outcome of each of these studies was summarised by Balls *et al.* (1999). No test was found capable of replacing the Draize rabbit eye test, but some of the assays

showed considerable promise as screens for ocular irritancy. For example, a validation study carried out in Germany led to the conclusion that the combined use of the HET-CAM and NRU tests could be used to identify severe eye irritants (chemicals with EU risk phrase R41). Subsequent analyses suggest that the endpoints of the of the 3T3 NRU and HETCAM tests can also be used to distinguish between non-irritants and irritants (i.e. chemicals with EU risk phrases R36 and R41), although the PMs, based on the data generated in the German validation study, have not yet been subjected to a formal validation (Worth & Cronin, 2001c).

The use of *in vitro* methods as screening tests is widespread in industry, since there is much confidence that a number of alternative tests do work in-house. However, it has proved impossible to establish this satisfactorily by conducting validation studies in which *in vitro* test results are compared with historical animal data. The main reason for this is the subjective scoring of tissue lesions in the eye in the Draize test, which provides variable estimates of eye irritancy. Other possible contributing reasons for the outcomes of recently completed validation studies are: a) the *in vitro* tests only partially modelled the complex *in vivo* eye irritation response, b) the protocols and PMs might have been insuffiently developed, and c) the choice of statistical approaches for analysing the data might not have been appropriate (Balls *et al.,* 1999).

Table 5.2 summarises those alternative methods which are currently the most developed and the most widely used. The Irritection system (formerly EYTEX™) is not recommended, due to the lack of a standardised protocol and poor *in vitro-in vivo* correlations.

It is generally considered that a battery of alternative tests is required for the assessment of eye irritation, since there are multiple mechanisms of eye irritation. For example, corneal opacity can result from the increased corneal hydration due to inhibition of the sodium-potassium ATPase pump in the corneal endothelium (Cejkova *et al.*, 1998). Such batteries should be based on *in vitro* tests that model different mechanisms and therefore give complementary results.

5.6.3 Structure-activity relationships for eye irritation

In a study by Cronin *et al.* (1994), the application of linear regression analysis to a data set of 23 physicochemical properties for 53 organic liquids led to the development of statistically significant QSARs (based on logP) for predicting the molar eye score (MES) of alcohols and acetates. The MES is the modified maximum average score (MMAS) corrected for the number of molecules applied to the rabbit eye.

Subsequently, Abraham *et al.* (1998a) used the solvatochromic parameters (molar refraction, polarisability, hydrogen bond acidity and basicity), which are derived from chromatography experiments, to model a data set comprising 38 of the 53 organic liquids previously analysed by Cronin *et al.* (1994). A QSAR for predicting the MES, based on these parameters and a vapour solubility parameter, had an r^2 value of 0.89.

A different approach to the prediction of the molar eye score was adopted by Kulkarni & Hopfinger (1999). In addition to using parameters based on solute properties, they also used a molecular dynamics method to generate intermolecular membrane-solute interaction properties. QSARs based on these properties were then derived by using a

genetic algorithm. A QSAR based on three parameters (two energy terms and a topological index) had an r^2 value of 0.92. The 16 chemicals used to derive the model were aliphatic and aromatic hydrocarbons, and aliphatic ketones, alcohols and acetates.

In addition to the derivation of regression models, attempts have also been made to develop classification models for eye irritation. For example, Cronin *et al.* (1994) applied linear discriminant analysis to the data set for 53 organic liquids, but found no linear combination of physicochemical properties capable of discriminating between irritant and non-irritant chemicals (as defined by EU classification criteria). However, in a PC plot based on all 23 variables, the irritant chemicals appeared to form an embedded cluster within the non-irritant chemicals. Similar findings were subsequently reported by Barratt (1995).

The phenomenon of embedded clustering of irritant chemicals was investigated further by Cronin (1996), this time by using the technique of cluster significance analysis (CSA) to determine whether the embedded clustering was statistically significant. Out of a total of 23 physicochemical descriptors, it was reported that the five most significant were logP, logP^2, the heat of formation, dipole moment, and a topological index. Subsequently, the method of "embedded cluster modelling" was developed (Worth & Cronin, 1999), which generates elliptic prediction models for the embedded clusters, and was applied to an eye irritation data set (Worth & Cronin, 2000).

The physicochemical determinants of eye irritation potential were also investigated by Rosenkranz *et al.* (1998). In comparison with non-irritant chemicals, these workers reported that irritant chemicals have significantly lower molecular weights, higher aqueous solubilities, lower logP values, and greater molecular orbital energy gaps (absolute hardness values). On the basis of the last-named observation, it was concluded that chemical reactivity does not appear to be a requirement for eye irritation. The authors also used the Multi-CASE expert system to identify biophores (substructures which occur with a significantly greater frequency in irritants than in non-irritants) and biophobes (substructures which occur significantly more frequently in non-irritants). The major structural determinants included primary, secondary and tertiary amine groups (i.e. basic groups), as well as carboxylate, organosulphate and sulphonate groups (i.e. acidic groups).

In contrast to the studies summarised above, which aimed to develop (Q)SARs for non-surfactant chemicals, an investigation by Patlewicz *et al.* (2000) focused on surfactants. In this study, neural network analysis indicated that the MAS of cationic surfactants is positively correlated with surfactant concentration and critical micelle concentration, and negatively correlated with with logP.

5.7 The ECVAM workshop on eye irritation and follow-up activities

In 1999, ECVAM organised a workshop on *Eye Irritation Testing: the Way Forward* (Balls *et al.*, 1999). Four parallel activities were suggested as a means of making progress toward the short-term reduction and refinement of animal use, and the long-term replacement of the Draize test: a) an evaluation of the reference standards (benchmarking) approach; b) a review of tiered testing strategies; c) further analyses

of the data obtained in previous validation and evaluation studies; and d) further research on the mechanisms of eye irritation. The following sections describe the progress made on these activities.

5.7.1 The reference standards approach

Following the outcomes of numerous validation and evaluation studies, it is now widely considered that the adoption of a reference standards (RS) approach could considerably improve the predictive abilities of *in vitro* tests for which no relevant and reliable *in vivo* benchmark exists. To investigate the applicability of the reference standards approach in the validation of *in vitro* tests, ECVAM established a Reference Standards Working Group, which decided that an initial evaluation of the benchmarking approach should be made by concentrating on eye irritancy as the toxicological endpoint.

Subsequently, an ECVAM Reference Standards Study was initiated in 1998, with the objective of assessing the feasibility of using reference standards to demonstrate the scientific validity of *in vitro* tests for eye irritation. Five *in vitro* methods were included in the study: a) the isolated chicken eye (ICE) test; b) the bovine corneal opacity and permeability (BCOP) test; c) the combined use of the HET-CAM and NRU test; d) EpiOcular^{TM} (a reconstituted human corneal epithelium); and e) the red blood cell (RBC) haemolysis test. Four groups of chemicals were tested: a) neutral organics (BCOP); b) alcohols and esters (HET-CAM/NRU; EpiOcular); c) surfactants (ICE; HET-CAM/NRU; RBC); and d) siloxanes (BCOP; ICE). Sets of five unknowns were tested against sets of five reference standards.

The outcome of the study revealed that the BCOP results were good for neutral organics, and that the RBC and NRU results were good for surfactants. Results for alcohols and esters were not as good, but data obtained with the HET-CAM/NRU test were better than data obtained with the EpiOcular test. Siloxanes were the most difficult group included in this study, because their physical properties made them difficult to handle consistently in some of the assays. The BCOP performed slightly more predictably than the ICE. This might be explained by the fact that, for this set, the *in vivo* score was presented not as the MMAS, but as a differently weighted PII. This might have reduced the *in vivo-in vitro* correlation with the ICE which generates a score intended to be a MMAS equivalent (Brantom *et al*., 2000).

The reference standards study led to the conclusions that:

1) The reference standards approach is feasible, but further investigations into its applicability are required.
2) The classifications of chemicals into different chemical classes in the study were too crude and too broad, and needed redefinition, since the assumption that reference standards would allow the testing of broad classes of chemicals with different mechanisms of effect proved false in this study.
3) The scientific validities of existing and new *in vitro* methods for eye irritation could be validated by using the reference standards approach.

The Management Team of the reference standards study recommended that:

1) With a view to re-defining chemical classes by identifying more sub-groups or classes of chemicals, the *in vivo* data should be analysed in terms of discrimination between different mechanisms of chemical action, by focusing on the effects of chemicals on different components of the eye, and on different time-scales of effect, regardless of apparent chemical similarity.

2) The limitations of each assay should be defined in terms of chemical class. The existing *in vitro* data should be re-examined in the light of the chemicals that succeeded or failed in each method, to identify how to better define groups of chemicals suitable for testing in each *in vitro* assay.

3) The domains of application of specific methods should be narrowed, and a decision tree for method selection should be developed.

4) A data base of reference standard chemicals should be established.

5.7.2 Tiered testing strategies

In the hazard identification of chemicals, which is different from the safety assessment of ingredients and mixtures of ingredients used in a wide range of industrial, pharmaceutical and consumer products, the purpose of testing is to classify eye irritation potential according to classification schemes defined by regulatory authorities.

In 1996, the OECD proposed a stepwise approach to hazard identification, which underwent revisions in 1998 and in 2001 (OECD, 2001). According to the OECD strategy, new chemicals can be classified as irritating to the eye on the basis of results from non-animal tests, including structure-activity relationships, physicochemical tests (such as the pH test), and *in vitro* tests. Testing in animals is only required as a last step to confirm negative results generated by the non-animal tests applied in earlier steps. The stepwise process is therefore intended to reduce and refine the use of animals, but the *in vivo* test is not replaced. In the EU, a tiered testing strategy for eye corrosion/irritation is being proposed for incorporation into Annex V of *Directive 67/548/EEC*. As with tiered testing for skin corrosion, this could be achieved during 2002, by means of the 29th Adaptation to Technical Progress (ATP) of the directive (Juan Riego-Sintes, personal communication).

Computer-based simulations of the OECD approach, using real data and prediction models, have shown that testing strategies based on the stepwise use of SARs, the pH test, an *in vitro* test, and the Draize test, can provide a satisfactory means of identifying eye irritants, while at the same time reducing and refining the use of rabbits (Worth & Fentem, 1999; Worth, 2000).

During the development and validation of a non-animal method intended as a screen in a stepwise testing strategy, it should be sufficient that the test can place chemicals into two or more categories of eye irritation potential, without generating too many false positive results. There is less concern about the generation of false negatives, because these would be identified by the animal test(s) carried out in the last step of the process. The revised OECD TG 405 (acute eye irritation/corrosion) is not in itself mandatory in OECD Member Countries, but it may become mandatory in an individual Member Country, if the regulatory authority of that country so decides.

The breakout group on *Acute Toxicity Testing Local: Skin and Eye* at the First International Symposium on Regulatory Testing and Animal Welfare, held in Ottawa, Canada, in June 2001, concluded that, although the use of screening tests is very common in industry, this is not necessarily identical to the use of non-animal tests in the OECD stepwise strategy, or in the Globally Harmonised System (GHS) for classification. *In vitro* alternatives are extensively used by industry for product selection, but in-house experiences are not being shared.

The development and implementation, in a regulatory testing framework, of appropriate testing strategies for eye irritation, which limit the use of the Draize rabbit test to the final step, are critically dependent on the availability of one or more scientifically validated *in vitro* tests for inclusion in the testing strategy. Therefore, in the short-term, the *in vitro* tests currently being used in-house must be demonstrated to be valid for the purposes to which they are being used, or new *in vitro* tests will need to be developed and validated.

5.7.3 Research on the mechanisms of eye irritation

The COLIPA Eye Irritation Task Force has developed a research programme, in collaboration with external academic partners and funded by the cosmetics industry (Bruner *et al.*, 1999). The aims of this programme are to: a) identify new *in vitro* endpoints that are more predictive of the *in vivo* response of the human eye than is the Draize test; and b) define the circumstances when it is appropriate to use different types of test systems, including simple epithelial cultures and three-dimensional cultures. Experiments will be conducted on human corneal cell cultures, rabbit cornea (from abattoirs), and human corneas which are unsuitable for transplantation, taking into account the experience obtain from earlier in-house research and from previous validation studies.

The International Life Sciences Institute (ILSI) Health and Environmental Sciences Institute (HESI) is also organising activities on alternative tests for eye irritation, such as the development of a mechanistically-based eye irritant classification scheme (Nussenblatt *et al.*, 1998). Further information is available from the following website: http://hesi.ilsi.org.

5.7.4 The acceptance of alternative methods by national regulatory agencies

ECVAM has conducted a survey to establish the uses for which alternative methods for eye irritation are formally accepted by regulatory authorities in the EU Member States. The outcome of the study revealed that:

1. In Belgium, the use of the BCOP test is accepted by the Pharmaceutical Commission for the identification of drug formulations that are eye irritants. For pesticides, the *in vivo* methods in Annex V of *Directive 67/548/EEC* are generally used, although positive results obtained in *in vitro* tests are also considered. For New Chemicals, only the methods defined in Annex V are accepted, whereas for Existing Chemicals, literature data are accepted, if the results are positive and considered to be adequate. To date, very few *in vitro* data have been received for the notification of new chemicals, and *in vitro* data have never been received for the registration of pesticides or biocides.

In Germany, the BgVV accepts *in vitro* data for the classification and labelling of severely eye irritating chemicals that are obtained with the HET-CAM test. For severe eye irritants, classifications may also be based on the results obtained in the BCOP, the IRE and the ICE test. In addition, chemicals that are corrosive to the skin or that are very acidic or basic (low or high pH in solution) can be classified as severely irritating to the eye, without additional *in vitro* or *in vivo* testing.

2. In The Netherlands, the ICE test is accepted by the regulatory authorities for the screening of severe eye irritants. The same applies to the BCOP and other tests based on isolated eyes, although these studies are rarely submitted to the Dutch Competent Authority. Although data from the HET-CAM test have not been submitted yet, this test will be accepted for the classification of severe eye irritancy provided that the chemical under investigation belongs to classes of chemicals for which the model has been shown relevant. Since the NRR assay and the agarose diffusion method have no clear prediction models for severe eye irritation, which could result in the overprediction or underestimation of R41 labelling, the Dutch Competent Authority has a preference for the other methods (isolated eyes, isolated corneas and HET-CAM). No *in vitro* method is accepted as a stand-alone test for the classification of irritating (R36) and non-irritating substances.

3. In the UK, the Health & Safety Executive (HSE) accepts the use of the IRE, BCOP and HET-CAM assays for the detection of severe eye irritants. These *in vitro* tests are acceptable for the testing of new and existing substances, as well as biocides. The position of the HSE is that where a positive result is obtained, the substance is considered a severe eye irritant (R41), and no further testing can be justified on animal welfare grounds. Where a negative result is obtained, an *in vivo* test may be required. The HSE prefers that tests are conducted in a GLP environment, and that the full test report and study protocol are provided in the regulatory submission.

4. The Irish Competent Authority has accepted *in vitro* data from the BCOP assay. *In vitro* data generated by other tests have not yet been submitted to the Authority.

5. In Finland, there is no official position on the general acceptance of methods other than those mentioned in Annex V of *Directive 67/548/EEC*. However, the reliability and relevance of non-validated *in vitro* methods are assessed on a case-by-case basis. In particular, if an *in vitro* test indicated that a chemical is irritating to the eye, the chemical could be classified as eye irritant on the basis of this test result. However, it is unlikely that a chemical would be classified as non-irritating on the basis of an *in vitro* method alone: a negative finding would have to be confirmed by using a method included in Annex V. The Finnish Competent Authority also stresses the importance of conducting all non-clinical health and environmental safety studies for regulatory purposes in accordance with GLP principles.

6. In France, the BCOP, HET-CAM, IRE and ICE are accepted for positive classification (R41with EU classification system or irritant cat.1 with GHS). If a method would be able to differentiate a moderate degree of irritation, a R36

classification could be envisaged. In this case, the French authorities would require extensive documentation of the method in a validation process. The neutral red release assay and the agarose diffusion method are accepted for the evaluation of cosmetic products.

7. The Greek Competent Authority only accepts official testing methods that are included in Annex V.

8. The Spanish Competent Authority stated that no *in vitro* data have yet been submitted to them. Its position is not to accept data from alternative methods, unless considerable experience with the methods exist or they have been validated and officially accepted. In very special cases however, they would be prepared to consider the *in vitro* data.

9. Denmark has not yet officially accepted any specific alternative method for the prediction of eye irritation, but positive results from alternative methods have already been used together with results from *in vivo* experiments.

10. In Sweden, the National Chemicals Inspectorate is prepared to accept *in vitro* data in combination with existing *in vivo* data, pH measurements and SAR data in an integrated testing strategy, such as the one proposed by the OECD (test strategy attached to OECD Test Guideline 405) on a case-by-case basis. However, the National Chemicals Inspectorate has not yet been faced with this situation. For cosmetics, the Medical Products Agency applies the SCCNFP "Notes of guidance for testing of cosmetic ingredients for their safety evaluation" (October 2000).

It is important to note that all the National Competent Authorities are waiting for the relevant bodies, such as ECVAM and other Commission services, and/or the OECD to respectively validate, accept and publish test guidelines on alternative methods, to facilitate acceptance of the *in vitro* methods.

5.8 Eye irritation: summary, conclusions and recommendations

Progress in the validation of alternative tests for eye irritation has been hampered by the lack of *in vivo* data of sufficient quality for use in validation studies. However, data from a number of tests, including the BCOP, HET-CAM, IRE and ICE tests, are already accepted by national regulatory authorities, on a case-by-case basis, for the identification of severe eye irritants. Furthermore, the tiered testing of eye irritation is accepted by OECD Member Countries as a means of reducing and refining the use of the Draize eye test in rabbits.

5.8.1 Short-term prospects

1. The acceptance during 2002 by the OECD Council of a revised TG 405 (acute eye irritation/corrosion), which includes as a supplement the OECD tiered testing strategy for eye irritation and corrosion.
2. Further investigations on the applicability of the RS approach (Section 5.7.1).
3. The validation of QSARs and/or expert system rulebases for eye irritation.

5.8.2 Recommendations

In the short-term, EU national regulatory authorities should consider harmonising their positions on the acceptance of the BCOP, HET-CAM, IRE, ICE, and other non-animal tests for eye irritation. It is therefore recommended that ECVAM should commission weight-of-evidence reviews on the use of the BCOP, HET-CAM, IRE and ICE tests, for eventual consideration by the ESAC.

In the medium-term and long-term, the application of genomics and/or proteomics technologies could lead to the identification of promising new endpoints for eye irritation, provided that attempts were made to relate the genomics/proteomics data to *in situ* eye irritation endpoints. Since more than one tissue of the eye is usually injured, the integrated use of different types of test should be explored. In particular, it would be useful to include a test that measures the time-course and degree of recovery of the eye. Time-to-recovery is a piece of information that affects classification and labelling.

Figure 5.1. A tiered testing strategy for dermal irritation and corrosion.

Step 1:	Consider existing human or animal data	⇒	classify as C or I, and stop assessment
	⇓		
	If NC and NI, or no information		
	⇓		
Step 2:	Apply a validated (Q)SAR	⇒	predict C and stop testing
	⇓		
	If NC or no (Q)SAR		
	⇓		
Step 3:	Apply a PM based on pH data	⇒	predict C and stop testing
	⇓		
	If no prediction or no pH data		
	⇓		
Step 4:	Apply a PM based on a validated *in vitro* test for skin corrosion	⇒	predict C and stop testing
	⇓		
	If NC		
	⇓		
Step 5:	Apply a PM based on a validated *in vitro* test for skin irritation	⇒	predict I and stop testing
	⇓		
	If NI or no test		
	⇓		
Step 6:	Perform a Draize skin test on a single rabbit	⇒	classify as C or NC
	⇓		
	If NC		
	⇓		
Step 7:	Perform a Draize skin test on a maximum of 3 rabbits	⇒	classify as I or NI

Figure 5.2. A tiered testing strategy for ocular irritation and corrosion.

Step 1: Consider evidence of skin corrosion or
      skin irritation potential
      (from dermal testing strategy; Figure 5.1) ⇒   classify as C or I, and stop
                                       assessment

      ⇓
      If NC and NI
      ⇓

Step 2: Apply a validated (Q)SAR   ⇒   predict as I and stop
      for eye irritation   testing
      ⇓
      If NI or no (Q)SAR
      ⇓

Step 3: Apply a PM based on a validated
      in vitro test for eye irritation   ⇒   predict as I and stop
                                         testing

      ⇓
      If NI or no test
      ⇓

Step 4: Perform a Draize eye test on
      a single rabbit   ⇒   classify as C
      ⇓
      If NI
      ⇓

Step 5: Perform a Draize eye test   ⇒   classify as I or NI
      on a maximum of 3 rabbits

Footnotes

In steps 2 and 3, QSARs and PMs that enable the identification of severe eye irritants
should be applied beofre models for the identification of non-severe eye irritants.

Table 5.1. **An overview of *in vitro* methods for skin corrosion and irritation.**

Method	Test System	Endpoint	Applicability	Formal Status
Skin corrosion				
Rat skin transcutaneous electrical resistance (TER) assay	excised rat skin	electrical resistance	general; additional dye-binding step for surfactants and solvents	validated and accepted by regulatory authorities for skin corrosion testing in the EU
EPISKIN™ human skin model (commercial system)	reconstructed human epidermal equivalent	cell viability (MTT reduction assay)	general; a few materials may interfere with MTT reduction	validated and accepted by regulatory authorities for skin corrosion testing in the EU
EpiDerm™ human skin model (commercial system)	reconstructed human epidermal equivalent	cell viability (MTT reduction assay)	general; a few materials may interfere with MTT reduction	validated and accepted by regulatory authorities for skin corrosion testing in the EU
CORROSITEX™ (commercial system)	reconstituted collagen matrix	colour or physical change in indicator "chemical detection system"	mainly acids, bases and derivatives	validated and endorsed (US and EU) for skin corrosion testing of acids, bases and their derivatives

Method	Test System	Endpoint	Applicability	Formal Status
Skin irritation				
EPISKIN™ human skin model (commercial system)	reconstructed human epidermal equivalent	cell viability (MTT reduction assay)	general; a few materials may interfere with MTT reduction	protocol modification and prevalidation (validation study under discussion)
EpiDerm™ human skin model (commercial system)	reconstructed human epidermal equivalent	cell viability (MTT reduction assay)	general; a few materials may interfere with MTT reduction	protocol modification and prevalidation (validation study under discussion)
Pig ear test	pig ear	trans-epidermal water loss (TEWL)	general	further development necessary
Mouse skin integrity function test (SIFT)	excised mouse skin	TEWL and electrical resistance	general; a few materials may interfere with either TEWL or ER determination	protocol modification and prevalidation (validation study under discussion)

Table 5.2. An overview of *in vitro* methods for eye irritation.

Method	Test System	Endpoint	Reference
Bovine corneal opacity and permeability (BCOP) test	Excised cornea from the bovine eye	Opacity and permeability of the cornea	Gautheron *et al*., 1992
Hen's egg test - chorioallantoic membrane (HET-CAM)	Hen's egg	Damage to the chicken chorioallantoic membrane	Lüpke, 1985
Chorioallantoic membrane – trypan blue staining (CAM-TBS)	Hen's egg	Damage to the chicken chorioallantoic membrane	Hagino *et al*., 1991
Isolated rabbit eye (IRE) test	Isolated rabbit eye	Corneal swelling, corneal opacity and fluorescein retention	Whittle *et al*. 1992
Isolated chicken eye (ICE) test	Isolated chicken eye	Corneal swelling, corneal opacity and fluorescein retention	Burton *et al*., 1981
Fluorescein leakage (FL) test	Madin-Darby Canine Kidney (MDCK) cells	Damage caused to the tight junctions in MDCK monolayers	Tchao, 1988
Neutral red uptake (NRU) test	3T3-L1 cells	Cell viability	Borenfreund & Puerner, 1985
Neutral red release (NRR) test	Rabbit corneal fibroblasts or mouse embryonic fibroblasts or normal human epidermal keratinocytes	Damage to the cell membrane	Reader *et al*., 1989
Red blood cell (RBC) haemolysis test	RBCs from calf blood samples	Damage to cytoplasmic membrane (haemolysis) in combination with damage of liberated cellular proteins (denaturation)	Pape & Hoppe, 1991
Agarose Diffusion Method	L929 mouse fibroblast cells	Cell death	Wallin *et al*., 1987
EpiOcular^{TM}	Reconstituted human corneal epithelium	Cell viability, release of inflammatory mediators, membrane permeability	Stern *et al*., 1998

6.   LOCAL TOXICITY: SENSITISATION

6.1 Introduction

Skin and respiratory sensitisation are considered separately from acute dermal and ocular toxicity, since the former endpoints are effects that may arise, following an initial exposure to a sensitiser (by any route of exposure and possibly via the systemic circulation), when a susceptible individual is subsequently exposed via the dermal or inhalation routes.

Skin sensitisation resulting in allergic contact dermatitis is an important occupational and product consumer health problem, leading to high costs to health care systems. It is therefore essential that chemicals and products of various kinds are evaluated for their skin sensitising potential. Until recently, the traditional tests for the identification of skin sensitisation hazards made use of guinea-pigs. However, with increasing understanding of the complex cellular and molecular events which induce skin sensitisation and elicit allergic contact dermatitis, it is becoming possible to consider alternative approaches for hazard identification and characterisation.

Current research on *in vitro* alternatives is mainly focused on one or more of the following properties of chemicals: inherent or potential protein reactivity, and the ability to penetrate the *stratum corneum* to reach the viable epidermal cell layers; ability to induce changes in, or in the responses of, cultured dendritic cells; and allergic stimulation of T lymphocyte activation.

The current status of these alternative approaches is reviewed below, and a stepwise testing strategy for the assessment of skin sensitisation potential is proposed.

6.2 Current status of alternative methods for skin sensitisation

6.2.1 The murine local lymph node assay (LLNA)

The LLNA is an *in vivo* mouse assay for skin sensitisation, which provides a reduction and refinement alternative to the traditional, guinea-pig methods, the guinea-pig maximisation test (GPMT) and the Buehler test. The LLNA also provides a more rapid quantitative and objective output, and uses only about half the number of animals required for the standard OECD protocols for the GPMT and the Buehler test. Furthermore, it does not require the use of Freund's complete adjuvant, intradermal injections of test substance, fur removal, occlusive dressings, the use of restraint, or the elicitation of an allergic skin reaction, all of which are features of the GPMT and/or the Buehler test.

The LLNA has undergone formal validation in the USA and Europe (NIH, 1999), has been endorsed as scientifically valid by the ESAC (ECVAM, 2000), and forms the basis of the draft OECD Test Guideline 429 (OECD, 2001), which will be considered by the OECD Council in 2002. In principle, the assay evaluates the extent to which a chemical stimulates the proliferation of lymphocytes in lymph nodes draining the site of chemical application. A chemical is regarded as a skin sensitiser, if it induces a

stimulation of proliferation which is ≥ 3 times that found in concurrent vehicle-treated controls (Gerberick *et al.*, 2001).

6.2.2 Knowledge-based computer systems

Relationships between the structure and biological properties of chemicals can be programmed into knowledge-based expert systems. One such expert system is DEREK ("Deductive Estimation of Risk from Existing Knowledge"; Sanderson & Earnshaw, 1991; Ridings *et al.*, 1996), which is under ongoing development by LHASA Ltd. (School of Chemistry, University of Leeds, UK). DEREK covers a variety of toxicological endpoints (e.g. mutagenicity, carcinogenicity, skin sensitisation), and is in widespread use in the chemical industry. Other expert system approaches to the prediction of skin sensitisation include the TOPKAT (Enslein *et al.*, 1997) and CASE systems (Graham *et al.*, 1996).

DEREK embodies both a controlling programme and a chemical rulebase (see also Section 9.4.1). The chemical rulebase consists of descriptions of molecular substructures called "structural alerts", which correlate with specific toxicological endpoints. The user communicates with DEREK by drawing the two-dimensional structure of the chemical under investigation on the screen. The rulebase is then searched against that structure, and any structural alert is highlighted, together with a message indicating the nature of the toxicological hazard.

The original skin sensitisation rulebase contained around 40 rules (Barratt *et al.*, 1994), which were derived from an historical database (Cronin & Basketter, 1994) containing data from guinea-pig maximisation tests on 135 chemicals that had been classified as skin sensitisers according to EU criteria. As a result of development of the system over the last seven years (e.g. Barratt & Langowski, 1999), the number of structural alerts for skin sensitisation currently stands at 59.

The following two-step strategy for the assessment of skin sensitisation potential is appropriate when using the DEREK rulebase:

1. The chemical is processed through the rulebase, to see if it has the potential to react with skin proteins either directly or, in some cases, after metabolism. If no structural alert is triggered, either the chemical does not possess the requisite reactivity, or its reactivity is outside the scope of the current knowledge base. In many cases, absence of chemical reactivity can be confirmed by inspection of the chemical structure. For chemicals that do not possess the appropriate chemical reactivity, no further computational evaluation is performed. However, an evaluation of possible metabolic activation of the compound is considered by expert input.

2. For chemicals or their metabolites that do possess the appropriate chemical reactivity, the second step is to assess their skin permeability/partition parameters. This initially involves using either empirical or calculated values for computation of the log octanol/water partition coefficient (logP) and/or to theoretically predict the log permeability coefficient (logKp) (Barratt, 1995). Molecular weight and melting point values are also used in the prediction algorithms to calculate logKp. Kp values can also be measured by using validated *in vitro* skin penetration

models. Using *in vitro* models also allows for the visualisation, as well as the quantification, of partitioning of compounds within the skin sub-structures.

6.2.3 In vitro tests

At present, no *in vitro* test for skin sensitisation has been validated, although several systems are in the course of development, based on an improved understanding of the biochemical and immunological mechanisms underlying the process. Among the key steps to be considered are protein binding, metabolism, and the cellular/immunological events leading to sensitisation. Cell-based assays investigated to date have included the use of human blood-derived dendritic cell cultures, Langerhans cells, keratinocyte cultures, human skin equivalents, and dendritic cell/T cell co-cultures (e.g. Kimber, 2000; Kimber *et al.*, 2001). These systems have been shown to express various mediators and/or markers of activation following exposure to chemical sensitisers, although not always with a reliable outcome (Ryan *et al.*, 2001).

A human reconstructed epidermis model, containing keratinocytes, melanocytes and Langerhans cells (Régnier *et al.*, 1998), is undergoing evaluation at L'Oréal (Rainer Schmidt, personal communication). Langerhans cells integrated into the reconstructed epidermis exhibit a reactivity to allergens similar to that found in the *in vivo* situation. The development of this system was supported by the EU BIOTECH programme (contract BIO 4 CT 960086).

Dendritic cell culture systems (Pichowski *et al.*, 2000; Tuschl *et al.*, 2000) represent a promising approach to the provision of a relevant test system. With the current increases in our knowledge of dendritic cell biology, coupled with the application of novel tools such as genomics and proteomics, it should be possible to gain new information on these systems, including a greater understanding of the interactions between chemical, dendritic cells and other cells within the skin, for future application in *in vitro* test systems. Possible hurdles in the development of a cell-based *in vitro* assay could be the initial haptenisation or protein binding events in sensitisation. A concerted effort to understand the mechanisms of chemical-protein interactions within the skin will provide invaluable information for the development of relevant *in vitro* test models. Investigations are currently under way, making good progress in both of these promising areas, although much more effort will be required before standardised, robust and reliable *in vitro* tests for chemical sensitisers will become available.

For the further development of *in vitro* tests for skin sensitisation, it is recommended that:

1) Basic research into the fundamental mechanisms of skin sensitisation should be continued, to identify opportunities for the development of novel, mechanistically-based *in vitro* assays.
2) *In vitro* systems for skin sensitisation should be biologically relevant (i.e. based on our understanding of the key mechanisms underlying sensitisation), while focusing initially on systems that model the initiation, rather than the promotion, of the immune response.

6.3 A tiered testing strategy for skin sensitisation

The proposal outlined below comprises a stepwise process (Figure 6.1) incorporating knowledge of physicochemical properties, knowledge of the relationship between chemical structure/reactivity and skin sensitisation potential, use of read-across, and finally, and when necessary, the LLNA, to confirm absence of sensitisation potential. Before the stepwise process is conducted for a chemical, an assessment should be made of its corrosivity potential. If the chemical is classified as either R34 (causes burns) or R35 (causes severe burns), skin sensitisation testing for hazard labelling purposes is not necessary. However, for risk assessment purposes, it might be necessary to conduct skin sensitisation testing at non-corrosive doses of a chemical labelled as R34 or R35.

Step 1: assessment of historical data
An initial screening of the available data will determine which chemicals already have adequate data relating to sensitisation potential. History-of-use data should also be assessed at this stage.

Step 2: assessment of physicochemical properties
Some chemicals are highly unlikely to possess skin sensitisation potential, due to their physicochemical properties, e.g. some inorganic salts (those elements which are sensitisers, e.g. nickel, chromium, are well-known). Some polymers with very high molecular weights (e.g. >5000) can also be ruled out on the basis of very low bioavailability, provided that there are no degradation products.

Step 3: screening of structures using the DEREK skin sensitisation rulebase
This process will highlight any structural features of a chemical which indicate a potential to react with skin proteins either directly or, in some cases, after metabolism in the skin (this is the reactivity component of skin sensitisation). However, DEREK does not possess rules for all possible cutaneous metabolic pathways, and an expert evaluation of likely metabolic activation should be conducted for molecules possessing no chemical structural alerts. Protein binding assays can also be used to assess the potential to react with skin protein.

Step 4: assessment of partition parameters
In order to behave as a skin sensitiser, a chemical must be able to partition through the skin and/or into an appropriate compartment for metabolism, and must possess the appropriate reactivity parameter (for covalent binding). Partition parameters can be calculated from physicochemical properties. Comparisons should be made by read-across to other chemicals possessing the same structural alerts, and for which sensitisation data are available. At this point, it may be possible to classify a chemical as a skin sensitiser on the basis of its reactivity potential and partition properties.

Step 5: in vitro assessment of skin sensitisation
In vitro studies should be performed for the classification of skin sensitisers, once such systems have been scientifically validated for this purpose.

Step 6: the murine local lymph node assay
In the absence of definite indications of skin sensitisation potential from steps 1 to 5, a murine LLNA should be performed according to OECD Test Guideline 429.

6.4 Skin sensitisation: summary, conclusions and recommendations

A number of methods, including QSAR mdoels and the DEREK skin sensitisation rulebase, reconstructed epidermis models, and dendritic cells are available, and could be used for priority setting. In cases where animal testing is required, the LLNA should be used in preference to the traditional guinea-pig tests, except for those classes of chemicals for which the LLNA is not considered to be appropriate. In addition to its traditional use for hazard identification, the LLNA can also be used for the determination of relative potency. Further work is needed before *in vitro* systems for skin sensitisation could be used for regulatory purposes (i.e. classification and labelling, and dose-response assessment).

6.4.1 Short-term prospects

1. The acceptance during 2002 by the OECD Council of draft OECD TG 429, which is based on the LLNA.
2. The validation of existing QSARs and/or expert system rulebases for skin sensitisation.

6.4.2 Medium-term prospects

1. The development of systems, based on our current and growing understanding of mechanisms involved in skin sensitisation. Progress is being made in the area of protein binding, with efforts being focused upon understanding the mechanisms of chemical-protein binding, which will ultimately yield the information required for the development of a standardised assay.
2. Human reconstructed epidermis models and human dendritic cell cultures represent promising cell-based approaches. Knowledge is growing in this area, and efforts are currently being focused on the identification of relevant cell phenotypes, and existing and novel endpoints for sensitisation.

6.4.3 Long-term prospects

The development and validation of new methods for skin sensitisation. The selection of a range of strong, moderate, weak and non-sensitisers should enable a determination of whether such methods could indeed provide a sufficient distinction between the strong, moderate and weak sensitisers.

6.4.4 Recommendation

The DEREK User Group should consider making more information available on the DEREK rulebase for skin sensitisation.

6.5 Current status of alternative methods for respiratory sensitisation

Certain chemicals are known to cause allergic sensitisation of the respiratory tract resulting in asthma and/or rhinitis. At present, no well-validated or widely accepted methods are available for the identification and characterisation of chemicals that have the potential to cause respiratory sensitisation. Originally, interest focused on the guinea-pig and the elicitation of pulmonary reactions in previously sensitised animals. More recently, alternative approaches have been described involving mice. The first

of these is the mouse IgE test, in which respiratory sensitising activity is measured as a function of the ability of a chemical to cause an increase in the total serum concentration of IgE (Warbrick *et al.*, 2002). The second approach is cytokine fingerprinting, in which chemicals with the ability to induce allergic sensitisation of the respiratory tract are identified on the basis of cytokine secretion patterns induced following *in vivo* exposure.

The current view is that the LLNA may be of some value in hazard identification. This is because the available evidence indicates that most, if not all, known chemical respiratory allergens elicit positive responses in the LLNA. The value of this is that chemicals which fail to induce positive LLNA responses can be regarded as lacking both contact and respiratory sensitising activity. Verification of this relationship will require further work.

When it is suspected that a chemical which is positive in a predictive test for skin sensitisation, such as the LLNA, might also be a respiratory allergen, further work should be undertaken to characterise the cytokine profile which results, since chemical contact allergens and chemical respiratory allergens induce divergent pathways of cytokine production (Dearman *et al.*, 1996; Dearman & Kimber, 2001).

Although there are as yet no well-characterised *in vitro* methods for predicting the potential of chemicals to cause respiratory sensitisation, the use of several systems for research purposes has been reported in the literature, including the use of human lung adenocarcinoma (A549) cells (Zhu *et al.*, 1999) and human bronchial epithelial (BEAS-2B) cells (Ramu *et al.*, 1996).

6.6 Respiratory sensitisation: summary, conclusions and recommendations

Respiratory sensitisation is an important endpoint in the context of occupational exposure to allergenic chemicals. There is no method for respiratory sensitisation in Annex V of *Directive 67/548/EEC*, and no *in vitro* test is sufficiently well-characterised for prevalidation. Therefore, further research is required, leading to the development of alternative methods in this area, with a view to providing *in vitro* methods for respiratory sensitisation for validation, acceptance and use in the long-term.

Figure 6.1. A tiered testing strategy for skin sensitisation.

Step 1:   Consider history-of-use data   ⇒   classify as a sensitiser,
and stop assessment

⇓

If non-sensitiser, or no information

⇓

Step 2:   Apply a validated PM based on   ⇒   classify as a non-sensitiser,
physicochemical properties   and stop assessment

⇓

If no decision

⇓

Steps 3 & 4: Assess potential reactivity   ⇒   classify as a sensitiser,
and partition parameters   and stop assessment
with validated (Q)SARs
and/or a validated expert system
rulebase

⇓

If non-sensitiser or no prediction

⇓

Step 5:   Apply a validated PM based on   ⇒   classify as a sensitiser,
in vitro data.   and stop assessment

⇓

If non-sensitiser or no prediction

⇓

Step 6:   Perform LLNA   ⇒   classify as a sensitiser or
non-sensitiser

7.   BIOKINETICS

7.1 Introduction

The systemic toxicity of a chemical to an organism depends on: a) the external dose of the chemical; b) the exposure route; c) the way in which the chemical is absorbed, distributed, metabolised and excreted from the organism (ADME/kinetics), which can be species-specific; and d) the intrinsic properties (biological activity/dynamics) of the chemical. These factors determine the critical concentration of a chemical at its target site(s) of action. Ideally, hazard and risk assessments should be based on the concentration-response curve(s) obtained at the target site(s), rather than the dose-response curve obtained for the whole organism.

This chapter describes the current status of non-animal methods for assessing barrier function (an important determinant of absorption and distribution) and metabolism. In addition, the use of biokinetic models is described, since these models provide a means of integrating information on absorption, distribution, metabolism and excretion, to provide a means of estimating the target tissue concentration of a chemical as a function of the external dose and time. Such information can be used in the design of *in vitro* and *in vivo* experiments for the testing of target organ and target system toxicity, thereby permitting the replacement of animal experiments or the reduction and/or refinement of animal use. It must be emphasised that many of the biologically-based parameters that need to be incorporated into biokinetic models can be obtained from *in vitro* experiments.

7.2 Barrier function

The following sections cover models of the skin, gastrointestinal and blood-brain barriers. Except in the case of the skin barrier, the permeability properties of the other barriers have mostly been studied in the context of pharmaceuticals rather than chemicals. However, it is important that knowledge about barrier properties obtained in the pharmaceuticals area is imported into the chemicals area.

7.2.1 Current status of QSARs for membrane permeability

In the QSAR modelling of membrane penetration, it has generally been assumed that membrane penetration occurs by passive diffusion across lipid bilayers. Only recently have attempts been made to model active transport, such as the role of P-glycoprotein in the oral absorption of drugs (van de Waterbeemd *et al.*, 2001). For a molecule to pass through a biological membrane, it must have certain molecular attributes. Permeability is related to hydrophobicity, and logP is the hydrophobicity descriptor employed in a number of QSAR models. Also, there is an upper limit on the size of the molecule that will pass through a membrane, which can be defined in terms of molecular weight, cross-sectional size, or another appropriate description of molecular bulk. The ability of a molecule to form hydrogen bonds is also known to be important, since hydrogen bond formation slows passage through membranes (Abraham *et al.*, 1995 & 1997).

Most QSARs for membrane permeability have been developed in the context of drug development, to predict the likely bioavailability of candidate drugs following administration (Cronin *et al.*, 2001). In principle, such QSARs are also applicable to the toxicity testing of chemicals, but they should be subjected to an independent evaluation of their mechanistic relevance and predictive performance, according to an approach similar to the one used in the validation of *in vitro* prediction models (Archer *et al.*, 1997; Worth *et al.*, 1998; Worth & Balls, 2001).

Table 7.1 provides a summary of the types of membrane permeability for which QSAR models are available, and some of the fundamental physicochemical descriptors that have been found to be useful for predicting various types of permeability. Unfortunately, there are no recent reviews that encompass the whole area of QSARs for membrane permeability. Numerous studies have been aimed at the development of QSARs for skin penetration, and these have recently been reviewed by Moss *et al.* (2001). An example is the following QSAR for skin penetration (Cronin *et al.*, 1999):

$$\log K_p = 0.77 \log P - 0.0103 \ MW - 2.3 \qquad \text{(Equation 7.1)}$$
$$n = 107, s = 0.394, r = 0.93, r^2 = 0.86$$

in which $\log K_p$ is the logarithm of the permeability coefficient across excised human skin, $\log P$ is the octanol-water partition coefficient, and MW is the molecular weight.

The QSAR models listed in Table 7.1 provide quantitative predictions of permeability. In certain circumstances, a qualitative prediction may be sufficient (i.e. a prediction of whether a compound is, or is not, likely to be absorbed). Lipinski *et al.* (1997) proposed the so-called "rule-of-five" for identifying drugs that would have poor oral absorption. This rule states that poor absorption is likely when any two of the following conditions are satisfied: a) molecular weight > 500; b) log P > 5.0; c) number of hydrogen bond donors > 5; and d) number of hydrogen bond acceptors > 10. A similar scheme has been proposed by Tice (2001) for agrochemicals. Similarities have been found between the rule-of-five and the properties associated with skin permeability (Cronin *et al.*, 1999). It is therefore likely that rules similar to the rule-of-five could be defined for a larger range of membrane types and for chemicals other than drugs.

Simple decision rules, such as the rule-of-five, have been successfully used to identify candidate drugs that are likely to be poorly absorbed and therefore to be of little therapeutic benefit. For toxicological purposes, it would be more appropriate to invert the rule-of-five. For example, a possible "inverse Lipinski" rule could state that good absorption is likely when any four of the following conditions are satisfied: a) molecular weight ≤ 500; b) log P ≤ 5.0; c) number of hydrogen bond donors ≤ 5; and d) number of hydrogen bond acceptors ≤ 10.

7.2.2 Current status of percutaneous absorption testing

The passage of a chemical across the skin is determined by the physicochemical properties of the chemical and the permeability properties of the *stratum corneum*, the outer layer of the epidermis, which consists of non-viable keratinocytes. Penetration across the skin and into the systemic circulation may result in lethal or sublethal toxic effects, and for this reason, the assessment of percutaneous absorption for regulatory

purposes is often carried out in the context of acute dermal toxicity testing. It should be noted, however, that the acute dermal toxicity test does not provide quantitative information on the amount of a chemical that crosses the skin barrier, but merely a gross indication of whether a sufficient amount crosses to cause systemic toxicity. Therefore, percutaneous absorption testing may also be carried out, to quantify the amount of a chemical that enters the systemic circulation.

A variety of methods are available for assessing skin penetration, including *in vivo* and *in vitro* methods, and QSAR models. *In vivo* methods, which provide biokinetic information on the whole-body distribution of the chemical, have involved many animal species, although the rat has probably been used most frequently (Howes *et al.*, 1996). The animal is killed at the end of the experiment, and the extent of percutaneous absorption is estimated from the known amount of chemical applied to the skin, and from the determination of the total amount excreted and of the amounts left on the skin and in the body. An *in vivo* method is described in the draft OECD Test Guideline 427 (OECD, 2000a). QSAR models are discussed in Section 7.2.1.

In vitro methods for skin absorption (Diembeck *et al.*, 1999) measure the diffusion of chemicals across excised (human or animal) skin, which may be of full or partial thickness. These methods can be based on the use of non-viable skin to measure diffusion only, or on the use of fresh, metabolically-active skin to assess diffusion, taking skin metabolism into account. An *in vitro* method is described in the draft OECD Test Guideline 428 (OECD, 2000b). There is good evidence that *in vitro* data are predictive of both human and animal data (Scott *et al.*, 1992; ECETOC, 1993). Nevertheless, for some classes of chemicals, poor *in vitro/in vivo* correlations have been reported (e.g. Bronaugh & Stewart, 1984). *In vitro* methods for percutaneous absorption were reviewed by Howes *et al.* (1996). A draft OECD Guidance Document for the Conduct of Skin Absorption Studies (OECD, 2000c) describes the circumstances in which the use of the *in vitro* method would be appropriate.

Following agreement by the OECD National Coordinators of the Test Guidelines Programme in May/June 2001, it is expected that revised versions of OECD Test Guidelines 427 and 428, and of the technical guidance document, will be adopted by an OECD Joint Meeting in 2002.

7.2.3 Current status of gastrointestinal barrier testing

The gastrointestinal barrier plays a dual role in the organism: a) to protect against toxic substances; and b) to transfer nutrients and xenobiotics (i.e. dietary products and drugs) from the lumen to the blood. The oral route of administration is the most common, so knowledge of the absorption and metabolism of a chemical at the intestinal mucosa level is important, since these processes affect the bioavailability of a chemical, which is defined as the fraction of an oral dose that reaches the systemic circulation (Stanley & Wilding, 2001).

Absorption across the intestinal lining includes passive diffusion between cell junctions (paracellular transport), passive diffusion across the cell membrane and cytoplasm (transcellular transport) or uptake via a transporter (active transport). Uptake transporters include the sodium-dependent bile transporter, peptide transporters, the glucose transporter, and organic anion transporters; export transporters include ATP-dependent export pumps (Ayrton & Morgan, 2001). P-

glycoprotein (P-gp) and multi-drug resistance proteins 1 and 2 (MRP1 and MRP2) represent well-characterised transporters in the apical membrane of the intestinal mucosal epithelium, which actively pump substrates back into the intestinal lumen after they have been absorbed into the intestinal epithelial cells (Ayrton & Morgan, 2001; Li, 2001; Li *et al.*, 2001).

The current status of *in vitro* models of the intestinal barrier has recently been reviewed by Le Ferrec *et al.* (2001). Table 7.2 provides a summary of the methods currently available for investigating the principal mechanisms of absorption, grouped as *in vivo*, *in situ* and *in vitro* methods (Hidalgo *et al.*, 1989; Ungell, 1997; Lennernas, 1998; Taipalensuu *et al.*, 2001).

The advantage of *in vivo* models is that these models integrate all the factors that can influence chemical partitioning. The disadvantage is that it is very difficult to separate the variables during absorption. The advantage of *in vitro* models over *in vivo* models is that it is possible to study the mechanisms of absorption *per se*, and by-pass stomach and liver metabolism.

ECVAM is funding a study on the development and refinement of an *in vitro* Caco-2 cell model of intestinal barrier function. This system consists of a cell monolayer across which a test chemical can be transferred. The chemical is usually applied as a solution in the donor compartment, and the amount transferred to a receptor compartment is measured. The goals of the ECVAM study are to standardise the use of Caco-2 cells for different purposes (absorption, metabolism and toxicity studies), and to improve and optimise the model with respect to the use of more-refined markers of intestinal function.

7.2.4 Current status of blood-brain barrier testing

The blood-brain barrier (BBB), formed by the brain-capillary endothelial cells and associated structures, and separates the brain parenchyma and the cerebrospinal fluid from the systemic circulation. It restricts the passage of chemicals from the circulating blood to the central nervous system. The brain capillaries are characterised by very close cell-cell contacts (tight junctions) that minimise the penetration of substances via the paracellular route. Chemicals must therefore penetrate via the transcellular route. Several transporters at the BBB play an important role in restricting the penetration of chemicals by removing them from the brain and transfering them to the systemic circulation. Consequently, some chemicals cannot achieve concentrations in the brain high enough to exert their potential biological activities (Stanness *et al.*, 1997; Kusuhara & Sugiyama, 2001; Cestelli *et al.*, 2001).

At present, there are no scientifically validated models of the BBB. The methods available can be classified into three groups: *in vivo* methods, primary cell cultures, and cell lines (Table 7.3). Several *in vitro* models have been designed to reproduce the physical and biochemical behaviour of the intact BBB; however, most of them lack some of the features of the *in vivo* barrier (Garberg, 1998; Veronese & Ehrlich, 1999; Foster & Roberts, 2000).

An ECVAM study on *in vitro* models for the BBB has recently been completed. The study included the characterisation and evaluation of: a) immortalised BBB derived endothelial cell lines (the rat-derived cell line, SV-ARBEC, and the mouse-derived

cell line, MBEC4); b) cell lines not derived from the BBB (MDCK, Caco-2 and ECV-C6 cells); and c) primary cells derived from the BBB (bovine brain endothelial cells and human brain endothelial cells). To evaluate these models, the *in vitro* data were compared with the *in vivo* data obtained in a mouse brain uptake assay. The aim of the study was to identify the *in vitro* models which best reflect the structural and functional characteristics of the BBB. It was found that the *in vitro-in vivo* correlations were low for all of the models evaluated, although the data obtained with the MDCK and Caco-2 cell lines correlated well with each other. One of the conclusions was that an MDCK cell line expressing the multi-drug resistance protein 1 (MDCKmdr-1) might be capable of distinguishing between chemicals that cross the BBB by passive diffusion and those that are substrates for active efflux. A workshop to consider the implications of the report on the study, and to propose further work, is to be held in Stockholm, Sweden, during 2002.

7.3 Barrier function: summary, conclusions and recommendations

Some QSARs for membrane permeability, such as those based on logP and MW, may be sufficiently predictive to identify chemicals that are likely to cross biological barriers by passive diffusion. These QSARs could be used for priority setting, but chemicals that are not predicted to undergo passive diffusion should not be neglected, since non-passive modes of membrane transport may also occur. All QSARs should eventually be validated in an independent manner, through an independent body such as ECVAM.

A variety of *in vitro* methods for percutaneous absorption are available and could be used for priority setting.

For gastrointestinal absorption, the Caco-2 culture model is considered to be sufficiently reproducible for use as a high-throughput screening system in the priority setting of chemicals. However, further work is needed to assess the predictive capacity of this system.

For assessing distribution across the BBB, a number of *in vitro* systems are under development. Further work is needed to validate these systems.

7.3.1 Short-term prospects

1. The acceptance by the OECD Council in 2002 of the draft OECD TG 428 (*in vitro* percutaneous absorption.
2. The further development of QSAR models for barrier function, including penetration across the skin, gastrointestinal barrier and BBB.
3. The further development of reconstituted human skin models for percutaneous absorption testing, to make their barrier properties similar to those found *in vivo*.
4. An evaluation of the feasibility of predicting bioavailability from *in vitro* data, and not just the fraction (percentage) absorbed, as currently performed by using Caco-2 model.
5. Further optimisation of a test protocol for the BBB involving primary endothelial cells co-cultured with primary astrocytes.

7.3.2 Medium-term prospects

1. The validation of QSARs for barrier function, including penetration across the skin, gastrointestinal barrier and BBB, following a preliminary assessment of their goodness-of-fit and mechanistic relevance.
2. The prevalidation of *in vitro* models of the gastrointestinal barrier and the BBB.

7.3.3 Recommendations for research and development

1. Further investigations on the expression of transport/efflux proteins in cell lines derived from the human gastrointestinal tract, and the influence of such transporters on absorption.
2. Research on the effects of anti-transport mechanisms (MDR, P-gp) and gut wall metabolism (CYP3A4) on bioavailability.
3. Further investigations on co-cultures consisting of cell lines with enterocytic markers and with mucus secretory functions, to increase understanding of the effects of mucus on the absorption rate.
4. The establishment of new cell models of the BBB that have characteristics more consistent with the *in vivo* situation, with an emphasis on the use of human cell lines.
5. The design and evaluation of a battery of *in vitro* assays, including, for example, measurements of protein binding and clearance, in addition to permeability, to predict brain distribution. Such a battery could include, for example, MDCK cells transfected with transporter proteins, such as mdr-1.

7.3.7 Recommendations

The ESAC should consider endorsing a statement on the applicability of *in vitro* methods for percutaneous absorption, assessed on the basis of a weight-of-evidence approach.

7.4 Xenobiotic metabolism

7.4.1 The scientific background to metabolism

Metabolism is the process by which a chemical is structurally changed in the body by enzymatic and/or non-enzymatic reactions. Information on the metabolism of a substance is important in the evaluation of its toxic potential. For example, the determination of metabolic stability can provide information on the potential for bioaccumulation. Information on the enzymes involved in metabolism can help to establish the importance of different clearance pathways, and can be useful in the prediction of clearance rates and the saturability of clearance processes. Metabolic studies can also be used to identify the tissue and cellular targets for any effect, and to establish the extent of inter-individual differences due to genetic polymorphisms, environmental factors or pathophysiological effects. The results of metabolic studies can also be useful when undertaking inter-species comparisons of the effects of chemicals.

A chemical that is absorbed orally is transported via the portal circulation to the liver, where it may be subjected to hepatic metabolism, followed by elimination in bile or via the kidneys. There is also the possibility of extra-hepatic metabolism. A typical drug metabolism pathway is the oxidation of the parent drug (phase I oxidation),

followed by conjugation of the oxidised moiety with highly polar molecules, such as glucose, sulphate, methionine, cysteine or glutathione (phase II conjugation). The key enzymes for phase I oxidation are the isoforms of the cytochrome P450 (CYP) family of enzymes (an updated list of which is available from: http://www.imm.ki.se/CYPalleles). The major human CYPs involved in chemical metabolism are CYP1A1, CYP1A2, CYP2A6, CYP3A4, CYP1B1, CYP2C9, CYP2C19, CYP2D6, and CYP2E1 (Clark, 1998; Soucek, 1999; Masimirembwa *et al.,* 1999; Ozawa *et al.,* 2000; Tanaka *et al.,* 2000; Riley *et al.,* 2001; Sheweita, 2001). The key phase II enzymes include N-acetyl transferases (NAT), UDP-dependent glucuronosyl transferase (UGT), phenol sulphotransferases (PST), oestrogen sulphotransferases (EST), and glutathione S-transferases (GST). The structures and functions of phase II enzymes have been reviewed elsewhere (Burchell *et al.*, 1998; Glatt, 2000; MacGregor *et al.*, 2001).

7.4.2 Current status of in vitro approaches for assessing metabolism

A number of *in vitro* systems are available for studying metabolism. These include: precision-cut tissue slices, subcellular fractions such as the microsomal fraction, primary cells in suspension, primary cells in culture, continuous cell lines, immortalised primary cells, and genetically-engineered cell lines.

To date, no *in vitro* methods for determining phase I or phase II biotransformation, or for evaluating metabolism-dependent toxicity, have been validated according to ECVAM's principles and procedures (Chapter 2). The current OECD Test Guideline 417 for assessing the toxicokinetic effects of chemicals (OECD, 1984) is based on *in vivo* studies. However, many studies provide support for the usefulness of *in vitro* methods for assessing metabolism and metabolism-dependent toxicity. From such studies, it is evident that there are enormous species differences in toxicokinetics; this is especially true of of the metabolic differences between between humans and rodents (Ozawa *et al.*, 2000). Therefore, there is a great need for human-based *in vitro* models that would offer better predictions of potential hazard to humans than could ever be obtained from laboratory animal studies (Coecke *et al.*, 1999, 2000, 2001; Doehmer *et al.*, 1999). The status of the currently available approaches for assessing metabolism is summarised in Table 7.4.

Recently, there has been success in relating the rate of *in vitro* metabolism of several compounds with the corresponding events *in vivo*. This represents a major step, since the value of *in vitro* metabolism systems has traditionally been considered to be purely qualitative in nature. The basis of this relationship is the use of the parameter, intrinsic clearance, which is a pure measure of enzyme activity toward a chemical *in vivo* and is not influenced by other physiological determinants of clearance, such as hepatic blood flow or drug binding. The *in vitro* equivalent of this parameter is the Vmax:Km ratio. The utility of *in vitro* intrinsic clearance as a predictor of *in vivo* intrinsic clearance, and therefore hepatic clearance and total body clearance, has been assessed with a database of 35 drug substrates for rat CYPs (Houston & Carlile, 1997). Other studies have demonstrated the usefulness of the intrinsic clearance approach for predicting kinetics in humans (Obach *et al.*, 1997; Ito *et al.*, 1998).

7.4.3 Current status of computer-based approaches for assessing metabolism

A number of computer-based expert systems for predicting metabolism and metabolism-dependent toxicity are undergoing development, and have been reviewed elsewhere (Dearden *et al.* 1995; Cronin, 2001).

An example is the METEOR system, a rule-based system for the prediction of metabolism which is being developed by LHASA (Green *et al.*, 1999). METEOR consists of a knowledge base of biotransformations that describe the metabolic reactions catalysed by specific enzymes, and that are related to specific substrates. The system can be linked with DEREK, to provide a means of assessing both metabolism and toxicity.

Two other rule-based expert systems, HazardExpert and MetabolExpert, are being developed and marketed by CompuDrug Chemistry Ltd. HazardExpert enables predictions to be made for a number of toxicological endpoints, taking into account factors such as species, dose, route and exposure duration. HazardExpert can be linked to MetabolExpert, which makes qualitative predictions of the metabolites of compounds. These predictions are made by using a rulebase consisting of molecular fragments. By linking HazardExpert with MetabolExpert, the user can obtain predictions of toxicity not only for the parent molecule, but also for the potential metabolites of the molecule.

Another approach to the prediction of metabolism (and metabolism-dependent toxicity) is provided by the Computer-Optimised Molecular Parametric Analysis of Chemical Toxicity (COMPACT) method of Lewis and coworkers (Lewis *et al.*, 1998). This method can be used to predict whether a molecule has the potential to act as a substrate for one or more of the cytochromes P450 (CYPs), or the ability to promote peroxisome proliferation. Oxidative metabolism by CYPs normally results in detoxification, although metabolism by CYP1, for example, may result in the formation of epoxides.

The COMPACT method is based upon the premise that there are certain structural requirements of a molecule that make it susceptible to oxidative metabolism. Firstly, molecules must be capable of binding to CYPs, due to their planar molecular shapes. Secondly, molecules that are capable of binding to a CYP must also be susceptible to chemical oxidation. The COMPACT approach is therefore based on two physicochemical descriptors, molecular planarity and electronic activation energy. Molecular planarity is a function of the cross-sectional area and molecular depth of the potential substrate, whereas the electronic activation energy is the difference between the energies of the highest occupied and lowest unoccupied molecular orbitals. A two-dimensional plot of molecular planarity and electronic activation energy for a series of molecules reveals that they can be divided into categories according to the particular CYP by which they are metabolised (Parke *et al.*, 1990).

The prediction rate obtained with the COMPACT system was found to be improved when its predictions were considered in combination with those generated by the HazardExpert system (Brown *et al.*, 1994), demonstrating the usefulness of a battery approach to prediction. The COMPACT approach has been extended to include the molecular (protein) modelling of the CYP enzymes themselves (Lewis *et al.*, 1999).

Another computational method for predicting potential metabolites is the META system, which is part of a suite of programs developed by Klopman and colleagues (Klopman *et al.*, 1994; Talafous *et al.*, 1994). The rules in the META rulebase were taken from the biochemical literature, rather than derived by a statistical approach.

7.4.4 A proposed strategy for assessing metabolism

So far, no consensus has been reached on the optimal test systems to be used for obtaining the information required. However, the most widely-accepted approach is to progress from general questions, involving the use of simple, inexpensive and less-specific *in vitro* models, such as human cell lines and human liver fractions, to mechanistic questions, which require technically more -demanding and more-complex models, such as cultures of human hepatocytes. In this way, the first tier acts as a preliminary screen for identifying the most important metabolic pathways, including metabolism-mediated toxic effects (Bull *et al.*, 2001), metabolic stability and enzyme inhibition (Coecke *et al.,* 2000). Recent developments in analytical chemistry (for example, in the use of mass spectrometry), have improved the value of *in vitro* systems for the determination of metabolic pathways and metabolic stability. Subsequent screens can then be used to study more-specific questions, such as the induction of biotransformation enzymes (Coecke *et al.,* 1999) or the occurrence of polymorphism-related effects (Coecke *et al.,* 2001).

Tier 1: Screening tests for metabolism
To identify the most important metabolic pathways, human liver fractions or genetically-engineered cell lines are incubated with the chemical, followed by quantitative analysis with techniques such as liquid chromatography (LC) and mass spectrometry (MS; Doehmer *et al.*, 1999; Ma *et al.,* 2001; Kassahun *et al.,* 2001).

To evaluate metabolism-mediated toxicity, *in vitro* cell culture assays can be used, which are based on primary monolayer cultures, cell lines, or genetically-engineered cell lines, which are expressing biotransformation enzymes (Schmalix *et al.,* 1996; Philip *et al.,* 1999, Coecke *et al.*, 2000. Bull *et al.,* 2001), and liver-derived cell lines re-expressing biotransformation enzymes (Coecke *et al.,* 1999). These *in vitro* tests can incorporate important toxicological endpoints, including cytotoxicity and genotoxicity (see Chapters 4 and 9, respectively). With all of these systems, it is necessary to take account of endogenous metabolism when interpreting the *in vitro* toxicity data.

To screen for metabolic stability, it is possible to use microsomes, human hepatocytes or genetically-engineered cell lines expressing human biotransformation enzymes. These approaches involve the incubation of the chemical with the cells or microsomes, followed by enzyme kinetic studies based on chromatographic procedures, such as high performance liquid chromatography (HPLC) and LC-MS (Linget & Vignaud, 1999). The exclusive use of microsomal preparations may give misleading results, because phase II enzymes, which are predominantly cytosolic enzymes and play a crucial role in the metabolic activation of chemicals, may be missing. In some cases, this problem can be solved by the exogenous addition of cofactors such as UDP-glucuronic acid (UDPGA) for glucuronidation, and 3'-phosphoadenosine-5'-phosphosulphate (PAPS) for sulphation (Swales & Utesch, 1998).

For the assessment of inhibitory interactions with CYP enzymes, the most popular approaches involve the use of either cultures of genetically-engineered cell lines containing only one specific human CYP isoform (Coecke *et al.,* 2000) or their microsomes (Miller *et al.*, 2000). The effect on the biotransformation of the substrate is dependent on the concentrations of both the inhibitor and the substrate, determined by the inhibition constant, K_i, of the inhibitor and the Michaëlis-Menten constant, K_m, of the substrate. The test compounds are incubated with the cultures or the microsomes in the presence of CYP substrates. Quantification can then be performed by appropriate fluorescence analysis, HPLC or LC-MS. In situations where more than one enzyme isoform is present, the use of complex *in vitro* models, such as human hepatocyte cultures (Li *et al.,* 1999), is advised, so that inhibition and metabolism can be evaluated at the same time.

In addition to enzyme inhibition by chemicals, there is also the possibility of immunoinhibition, in which polyclonal and monoclonal antibodies may competitively inhibit biotransformation enzymes (Rodrigues, 1999).

Tier 2: Induction of biotransformation enzymes

Xenobiotics taken up by the human body may result in the *de novo* synthesis of enzyme molecules (including phase I and phase II biotransformation enzymes) as a result of increased transcription of the respective gene. An increase in enzyme activity may also be observed as a result of enzyme stabilisation (Coecke *et al.,* 1999).

Various *in vitro* metabolically-competent models have been proposed for detecting the induction of CYPs, including precision-cut liver slices, short-term and long-term hepatocyte cultures, liver-derived cell lines expressing or re-expressing biotransformation enzymes, and highly differentiated human cell lines (Coecke *et al.*, 1999, Gomez-Lechon *et al.,* 2001). For detecting enzyme induction, these *in vitro* methods involve the use of endpoints such as: a) the activity of the biotransformation enzymes (and, if appropriate, their individual isoenzymes); b) protein levels, by using techniques such as immunoblotting or HPLC; and c) mRNA levels, by using Northern blotting, e.g. via the nuclease protection assay or the reverse transcriptase polymerase chain reaction (Coecke *et al.,* 1999).

Although there are technical difficulties associated with the cryopreservation of hepatocytes, a number of cryopreserved hepatocyte cultures have been used for enzyme induction studies (Hengstler *et al.*, 2000). To date, no prevalidation study has been carried out on methods for identifying the capacity of compounds to induce biotransformation enzymes. However, ECVAM will shortly initiate a prevalidation study (Coecke *et al.*, 1999) on the use of human hepatocyte sandwich cultures (LeCluyse, 2001).

Tier 3: Models for evaluating polymorphic effects on metabolism

Amino acid substitution or deletion can result in reductions in, or even loss of, the activities of phase I or phase II biotransformation enzymes. Certain individuals exhibit a severely compromised ability to metabolise chemicals that are specific substrates of these polymorphic enzymes, which can lead to serious toxic side-effects (Wolf & Smith, 1999).

The acetyltransferases, NAT1 and NAT2, represent the best-understood polymorphic enzymes. Other phase II enzymes, such as sulphotransferase (ST), UGT, and GST, are known to contain a variety of polymorphic variants, but their functional genetic diversity is not as well understood (Burchell *et al.*, 1998, 2000; Eaton & Bammler, 1999; MacGregor *et al.*, 2001; Glatt, 2000).

In a review by Friedberg *et al.* (1999), a comparison was made of insect, bacterial, yeast, and mammalian metabolism models. In general, the catalytic properties of CYPs in the various models were rather similar.

A recent and promising innovation is the use of established cell lines, such as V79 cells, which have been genetically engineered to express selected genetic variants of human CYP enzymes, so that such polymorphic effects can be assessed (Coecke *et al.*, 2001).

7.4.5 General recommendations for the use of in vitro metabolism tests

It is recommended that the *in vitro* metabolism tests used to implement the future Chemicals Policy should provide a simple, rapid and inexpensive means of detecting a broad array of chemically-mediated metabolic effects. The models should be able to provide information on: a) the most important metabolic pathways; b) the specific human enzyme isoforms involved in the metabolism; c) the metabolic stability of the chemical; d) the structures of the metabolites; e) metabolism-dependent toxification and detoxification; and f) the regulation of biotransformation enzymes, e.g. their capacity to be induced or inhibited.

A variety of high-throughput *in vitro* metabolism methods are available for use in tier one of the metabolism strategy (for a review, see White, 2000). ECVAM, in consultation with experts in the field of metabolism, has plans to initiate a number of prevalidation and validation studies in the future.

7.5 Metabolism: summary, conclusions and recommendations

A wide variety of *in vitro* tests (referred to as tier 1 tests in this report) are available for identifying metabolic pathways, metabolism-mediated toxic effects, metabolic stability and enzyme inhibition, and these could be used immediately to obtain mechanistic information. Other tests are available for assessing enzyme induction (tier 2) and polymorphic effects (tier 3). An important question for the implementation of the EU Chemicals Policy will be whether tier 2 and tier 3 tests are necessary, and if so, at what tonnage level of production/importation.

7.5.1 Short-term prospects

1. The prevalidation of *in vitro* tier 1 metabolism tests.
2. The prevalidation of *in vitro* tests for induction based on human hepatocyte cultures as tier 2 systems.
3. The further improvement of techniques for the cryopreservation of human hepatocytes.
4. The further development of models for evaluating polymorphic effects on metabolism.
5. The further development of computer-based systems for predicting metabolism.

7.5.2 Medium-term prospects

1. The validation of *in vitro* tier 1 metabolism tests.
2. The validation of *in vitro* tests for induction based on human hepatocyte cultures s tier 2 systems.
3. The prevalidation of models for evaluating polymorphic effects on metabolism.

7.5.3 Long-term prospects

1. The validation of computer-based systems for predicting metabolism from chemical structure.
2. The validation of models for evaluating polymorphic effects on metabolism.

7.6 Biokinetic modelling

Biokinetic models are mathematical models based on differential equations that describe the absorption, distribution, metabolism and elimination of chemicals as a function of dose and time. Biokinetic models can be divided into two main types: a) data-based compartmental models; and b) physiologically-based compartmental models, which are also known as physiologically-based biokinetic (PBBK) models. PBBK models are based on the known anatomy and physiology of the organism. Compartments correspond to relevant anatomical structures, such as the liver and kidney, or to tissue types, such as fat and muscle (Andersen, 1991; Clewell & Andersen, 1986). Over the last decade, the development of PBBK models has been facilitated by the increasing availability of computer-based techniques that can simultaneously perform the numerical solution of the sets of differential equations that characterise biokinetic processes. The use of PBBK models is expected to contribute to the replacement, reduction and refinement of animal studies, by enabling the design of studies to be optimised through the identification of critical kinetic parameters and time-frames.

The main advantage of PBBK models over classical data-based compartmental models is their ability to estimate kinetic behaviour on the basis of a predetermined model structure and independently-derived model parameters, thereby avoiding the need to adjust the parameters to optimise the fit of the model. PBBK models are based on two kinds of parameters: a) species-specific anatomical and physiological parameters, such as the pulmonary ventilation rate, for which data are generally available in the literature (Brown *et al.*, 1994); and b) compound-specific parameters, such as tissue-blood PCs, and kinetic parameters, such as the Michaelis-Menten constants, Vmax and Km. Compound-specific parameters have traditionally been obtained from kinetic studies performed *in vivo*. However, there is now considerable scope for deriving such parameters from *in vitro* studies studies (Blaauboer *et al.*, 2000), which should therefore lead to a reduction in the number of animal studies conducted for biokinetic modelling.

When biologically-based parameters are assigned to each of the specific kinetic processes, it becomes clear that once the compound is taken up in the systemic circulation, the distribution, metabolism and excretion of a chemical are independent of the exposure route. Thus, it is possible to extrapolate from one exposure route to another by adjusting the parameters for the relevant uptake process only, since the

mathematical descriptions of the distribution and elimination processes remain the same. In addition to route-to-route extrapolation, PBBK models also permit extrapolations of dose and species beyond the conditions of laboratory studies.

The distribution of a compound throughout the body is described by tissue volumes, blood flow rates, tissue-blood partition coefficients (PCs), and, where appropriate, by the kinetic parameters of active transport processes. Tissue volumes, tissue-blood PCs and blood flow rates can be determined experimentally or may be available from the literature.

Tissue-air PCs of volatile compounds can be measured *in vitro* by incubation of the compound with a homogenate of the respective tissue in buffered saline. The equilibrium distribution of the compound is then measured by gas-chromatographic analysis of the air present in the headspace of the vial containing the homogenate/buffer mixture. Tissue-blood PCs can be calculated by dividing the tissue-air PCs by the respective blood-air PCs. This method has been applied in many studies on volatile hydrocarbons, with both human and rat blood, and liver, lung, kidney, fat, muscle, and brain homogenates (Sato & Nakjima, 1979; Perbellini *et al.*, 1985; Fiserova-Bergerova & Diaz, 1986). This principle has been extended to determine the tissue-blood PCs of non-volatile compounds by the use of a number of *in vitro* methods (Pacifici & Viani, 1992; Jepson *et al.*, 1994). However, these techniques are time-consuming and often depend on the availability of a radiolabelled form of the chemical under consideration.

In addition to these *in vitro* techniques, both descriptive and mechanistic algorithms for the estimation of blood-air and tissue-blood PCs have been reported (DeJongh *et al.*, 1997 & 1998). Such methods attempt to estimate the biological partitioning process as a function of physicochemical parameters such as the aqueous and/or lipid solubilities of the chemical.

The distribution rate of a compound to the brain can be described mathematically as a function of input into the brain from the arterial blood stream and outflow from the venous side. A venous equilibration model (Andersen, 1991) makes the assumption that the concentration of a chemical in the venous blood leaving the brain is always in equilibrium with that in the brain tissue. This model is expected to be valid for small nonpolar, lipophilic chemicals, as well as hydrophilic ones. Thus, the perfusion rate of the brain dominates the diffusion rate through the blood-brain barrier as the rate-limiting step in brain distribution. Depending on its molecular structure, a chemical may exhibit a diffusion-limited brain uptake *in vivo*, in which case a compound-specific parameter is needed to express its diffusion rate. In the case of polar and/or ionisable compounds, an energy-dependent, saturable uptake rate may be required to describe its interaction with the BBB. *In vitro* models of the BBB are available (Franke *et al.*, 2000), and are described in more detail above (Section 7.2.5).

7.7 Biokinetic modelling: summary, conclusions and recommendations

PBBK models describe the processes of absorption, distribution, metabolism and excretion by integrating physicochemical and *in vitro* data. They can be used to determine target organ/system doses and to extrapolate between routes of exposure and between species. To date, most of the models have been developed for specific

chemicals, so there is a need to develop and validate generic models, applicable to broad groups of chemicals.

7.7.1 Short-term prospects

1. For the prediction of target organ/tissue distribution, a user-friendly software package is required, so that the user can enter a SMILES code or CAS number for a given chemical, then obtain a qualitative indication of the major target organs and tissues. This information could then be used to indicate the *in vitro* tests that could be performed for a chemical known to enter the body by the dermal route of exposure.
2. The further development of algorithms for predicting *in vivo* metabolic clearance from *in vitro* data.

7.7.2 Medium-term prospects

1. The validation of algorithms incorporated into the biokinetic software package.
2. The validation of algorithms for predicting *in vivo* metabolic clearance.

Table 7.1. An overview of QSARs for membrane permeability.

Membrane / barrier	Significant physicochemical descriptors	Reference
Cell membrane	Linear-free-energy-relationship descriptors, in particular hydrogen bond acidity	Platts, 2000
Caco-2 cells	Polar surface area, hydrophobicity Hydrophobicity, hydrogen bonding ability Polar surface area	Egan *et al.*, 2000 Österberg & Norinder, 2000 Stenberg *et al.*, 2001
Gastrointestinal tract	Polar surface area, hydrophobicity Hydrophobicity, hydrogen bonding ability Linear-free-energy-relationship descriptors	Egan *et al.*, 2000 Österberg & Norinder, 2000 Zhao et al., 2001
Skin	Hydrogen bonding ability Hydrophobicity, molecular size	Abraham, 1997 Barratt, 1995 Cronin, 1999 Kirchner *et al.*, 1997
Cornea	Hydrophobicity and molecular size	Worth & Cronin, 2000
Blood-Brain Barrier	Hydrophobicity, hydrogen bonding ability, molecular size Hydrophobicity, hydrogen bonding ability, molecular size Polar surface area, hydrophobicity, hydrogen bonding ability Hydrophobicity, hydrogen bonding ability, molecular size Hydrophobicity, hydrogen bonding ability	Clark & Pickett, 2000 Duffy *et al.*, 1999 Feher *et al.*, 2000 van de Waterbeemd, 2001 Österberg & Norinder, 2000

Membrane / barrier	Significant physicochemical descriptors	Reference
Alveolar membrane of the mammary gland	Many, including hydrophobicity, acid dissociation constant and molecular size	Agatonovic-Kustrin, 2000
Polydimethylsiloxane (Silastic) Artificial Membrane	Hydrogen bonding ability	Cronin *et al*, 1998, 2001

Table 7.2.   An overview of gastrointestinal barrier models.

Models	Advantages	Limitations
In vivo Rat model	Integrates all factors that can influence drug dissolution; It reflects the human situation in respect to paracellular space; Measures bioavailability.	Animal model; Differences regarding metabolism are quite common; Difficult to dissect variables in the absorption process.
In vivo Administration of specially designed capsules to human volunteers.	Allows regional absorption studies; Human model; Regulatory authorities are suggesting the importance of this technique to develop sustained release products.	Not appropriated when there are limited toxicological data available on the compound.
In situ Perfused rat preparations	Integrates passage and metabolism aspects; All physiological factors that affect passage are present; Studies of direct effects of the drug on intestinal absorption are possible; Allows preliminary screening.	Animal model; The increase of luminal hydrostatic pressure during experiment can influence intestinal permeability; Usually studies disappearance of compounds rather than appearance of compounds.
In situ Balloon technique in humans	It is a reference technique; Many physiological factors that influence passage are present; Allows studies in humans; Possible to study compounds secretion into the intestinal lumen.	A difficult technique; Expensive technique. Not used in development and not used routinely; Studies disappearance of compounds and it is less sensitive for studies on low permeable compounds; Introduce non-physiological conditions in the part of the intestine studied.
In vitro **organotypic** Gut sac of the rat small intestine	All cell types and mucus layer are present; Relatively fast and inexpensive technique; Useful for mechanism of absorption and formulation studies; Measures absorption at different sites in the small intestine.	Animal model; Not a perfused model; The drug must cross the whole intestinal wall; Low tissue viability (up to 2 hours).

In vitro **organotypic** Isolated and perfused segments	Possibility to study specific functions at the organ level; Physiological cell-to-cell contact and native intracellular matrix are present.	Short duration of experiments because of cell alterations.
In vitro **organotypic** Ussing chambers	A human and animal model; Drug absorption and passage at specific intestinal sites are possible; The test drug can be added on the apical and/or the basolateral side; Metabolisms studies are possible; Useful to study substances having local-pharmacological and transported-mediated effects; Possible to study electrophysiological parameters of the intestinal barrier.	Cell viability is limited; Availability of human material is limited; Not used for screening.
In vitro **cell model** MDCK cells	Fast and simple method; Can be used for screening testing; Can be used for measurement of passive diffusion;	Not an intestinal model; It is not human origin; Physiological factors that influence passage are not present (mucus, bile salts, cholesterol); A static model; Model with only one cell type.
In vitro **cell model** Caco-2 cells	Human cells; Display the majority of morphological, transport and permeability features of differentiated intestinal cells; Relatively fast and simple method; A flexible model; Allows mechanistic transport studies; Drug can be added at the apical or basolateral side;	Physiological factors that could influence passage are not present (mucus, bile salts, cholesterol); A static model; Cells have tumoral origin; Model with only one cell type; Low levels of CYP3A4;

	Allows drug screening testing; Predicts passive human passage; Useful for ranking drugs according to permeability; using reference compounds it is also possible to predict % fraction absorbed for passively transported compounds; Already used during drug discovery; Well establish model; Good availability of "reference" data.	
In vitro **cell model** TC7 cells (derived from Caco-2 cell line)	Express higher levels of CYP3A4; Grow faster than Caco-2 cells;	Physiological factors that could influence passage are not present (mucus, bile salts, cholesterol); A static model; Cells have tumoral origin; Model with only one cell type;.
In vitro **cell model** HT29-18-C1 cells	Less transepithelial electrical resistance than Caco-2; Paracellular transport more similar to small intestine (but still lower)	Not well established
In vitro **cell model** HT29-MTX/Caco-2 co-cultures	Cell model containing mucus-secreting cells	Not well established
In vitro **cell model** 2/4/A1 (conditionally immortalized epithelial cell line)	High permeability of the paracellular space; Gives similar correlation with fraction absorbed as human jejunal permeability.	Rat cell line; Not well established
In vitro **cell model** Primary isolated epithelial cells		Difficult to culture; Limited viability; Loss of *in vivo* anatomical and biochemical features.

Table 7.3.   **An overview of blood-brain barrier models.**

Models	Advantages	Disadvantages
In vivo Tissue distribution assay	Studies of brain uptake *in vivo*, either after single dose or steady state.	Animal model; Several animals needed; Number dependent on number of time point used; Not human
In vivo Microdialysis	Studies of brain uptake *in vivo*, either after single dose or steady state; Single animal used to study kinetics (continuously sampling from one animal over time); Possible to study kinetics simultaneously in several tissues, in the same animal.	Animal model; Low throughput; Not easy to set up or use.
In vivo Positron emission tomography		
In situ Perfused head		Technically demanding

Primary co-cultures Bovine brain capillary endothelial cells co-cultured with astrocytic cells	Well established; Best model at present; Morphology similar to *in vivo* situation, with specific markers and transporters retainedd; Provide important information about the functional properties of the BBB and the importance of cell-cell interactions in establishing this barrier.	Scarcity of the endothelial cells from brain microvessels; The expression of BBB characteristics decreases with time in culture; Variations in the purity of the culture markers makes it difficult to compare experiments; Lower trans-epithelial resistance and higher paracellular leakage, compared to *in vivo* situation; Culture conditions and cell preparations are not standardised among laboratories.
Immortalised cell lines • SV-ARBEC (rat brain microvascular endothelial cells) • MBEC4 (mouse brain capillary endothelial cell line) • REB4 (rat brain endothelial cell line) • REC-T1 (rat brain microvascular cell line) • SV-BEC (bovine brain) • CR3 (rat brain endothelial cell line)	Retain several endothelial markers; Easier to keep in culture and transfer between laboratories as compared to primary cells.	More leaky across paracellular pathway as compared to *in vivo* BBB and primary cultures; Variable expression of transport proteins, compared to *in vivo* BBB.
Immortalised cell lines RBE4.B co-cultured with cortical neuronal cells.	This co-culture system induces the characteristic permeability limitations of a functional BBB.	

Passaged or cloned cells From brain endothelial cells	Cells could be subcultured for up to 50 generations; Large quantities of brain endothelial cells are obtained; Still retain endothelial markers and some BBB markers.	Major differences are observed between different clones with regard to tightness of monolayers and other properties.
Spontaneously immortalised cell lines HUVEC-304 with and without co-culturing (C6 glioma cells)	Easy to obtain; Cultured in astrocyte-conditioned medium gain BBB properties.	Not a BBB derived cell line (expression of transporters and receptors likely to be different).
Established cell lines MDCK	Have been shown to generate data on permeability correlating to that obtained with primary bovine brain endothelial cells, as well as *in vivo* data.	Not a human cell line; Not a BBB derived cell line (expression of transporters and receptors likely to be different).
MDCK mdr1	Used in combination with MDCK wild type cells to identify P-gp substrates, which are likely to be restricted as regarding passage across BBB.	Not a human cell line; Not a BBB derived cell line (expression of transporters and receptors likely to be different).
Endothelial cells from pial microvessels	Easier to obtain than brain endothelial cells.	The type of tight junctions and the distribution of endothelia barrier antigen are different than brain endothelial cells.

Table 7.4. **An overview of *in vitro* methods for assessing metabolism.**

Strategies	Test systems	Endpoints	Applicability	Formal Status
First tier	Microsomes from human hepatocytes or from genetically engineered cell lines expressing human genes	Identification of metabolite formation by LC-MS/MS	Most important metabolic pathways	Prevalidation studies to be initiated
	Cell lines, primary monolayer cultures, genetically engineered cell lines expressing human genes	Cell morphology, viability, membrane damage, liver-specific endpoints, genotoxic endpoints	Metabolism-mediated toxic effects	
	Microsomes, human hepatocytes, genetically engineered cell lines expressing human genes	Quantification by LC-MS of the amount of the parent compound that remains after metabolism	Metabolic stability	
	Microsomes, human hepatocytes, genetically engineered cell lines expressing human genes	Quantification by LC-MS, HPLC or fluorescense	Inhibition	
Second tier	Short-term and long-term hepatocyte cultures (e.g. human hepatocyte sandwich cultures), precision-cut liver slices and liver-derived cell lines expressing or re-expressing biotransformation enzymes, highly differentiated human cell lines	Assaying the activity of the biotransformation enzymes (and, if appropriate, their individual isoenzymes); quantifying the protein level by using techniques such as immunoblotting or HPLC; and quantifying the mRNA levels by using Northern blotting, the nuclease protection assay or the reverse transcriptase polymerase chain reaction	Induction	Prevalidation study to be initiated
Third tier	Genetically engineered cell lines expressing human genes	Quantification by LC-MS	Polymorphism	Prevalidation study to be initiated

8.   TARGET ORGAN AND TARGET SYSTEM TOXICITY

8.1 Introduction

The detection of specific target organ and target organ system toxicity is an important aspect of toxicological testing, and the conventional test procedures involve chronic, repeat-dose *in vivo* testing, often in a non-rodent species as well as in a rodent species. This kind of predictive testing represents an enormous challenge to those who, for a variety of scientific and ethical reasons, would like to see animal testing replaced by non-animal tests and testing strategies. It will be worthwhile to put considerable skill and effort into attempting to meet this challenge, not least because, as in the case of carcinogenicity testing (Chapter 9) and reproductive toxicity testing (Chapter 10), inter-species differences limit the usefulness of animal studies for predicting long-term target organ and target system effects in humans. It should also be borne in mind that strategies for assessing target organ and system toxicity need to take into consideration the biokinetic considerations of absorption, distribution, metabolism and excretion (Chapter 7).

During the past decade, several *in vitro* systems have been in the course of development for detecting a number of target organ and target system effects (Spielmann *et al.*, 1998), and it will be vital, if only for reasons of the costs incurred and the time required, that strategies be developed for the selective and sequential use of a widening range of optional and complementary approaches. In this chapter, preliminary approaches to establishing *in vitro* systems for long-term, repeat-dose testing are discussed, and an example will be given of developments in relation to a target organ (the kidney) and to a target system (the nervous system).

8.2 Current status of repeat-dose toxicity testing

Chronic toxicity is aa consequence of the persistent or progressively deteriorating dysfunction of cells, organs or multiple organ systems, resulting from long-term exposure to a chemical. In animals (usally rodents), short-term repeat-dose studies last from 14 to 28 days (OECD TG 407 [oral toxicity], TG 410 [dermal toxicity] and TG 412 [inhalation toxicity]). The highest dose administered is designed to cause some toxicity, but not lethality. Upon completion of the test, a whole host of clinical and histological evaluations are recorded, including experimental observations and whole body and individual organ analyses. Other subchronic toxicity studies include the 90-day study (OECD TG 408), and the combined repeat-dose toxicity and reproductive toxicity screening test.

During long-term exposure, some compounds can induce adaptive processes in their target cells. Changes in gene expression at the transcriptional, translational or post-translational levels can result from the induction or inhibition of enzyme systems, or by up-regulation or down-regulation of receptor-ligand interactions, thereby changing the functions of intracellular signal cascades. For instance, the increase of metallothionein expression following exposure to low doses of metals permits a higher toxic dose to be tolerated (Miles *et al.*, 2000).

Another important parameter to take into consideration is the ability to recover from the toxic insult. Some *in vitro* studies on recovery and repeat dose testing have been published; however, the relevance of measuring reversibility *in vitro* and the interpretation of the *in vitro* results in terms of the *in vivo* situation still needs to be established (Combes, 1999).

There have been a few attempts to study the long-term exposure of cells or tissue cultures to chemicals *in vitro*. In most of these studies, non-organ specific parameters (e.g. proliferation, protein synthesis, and mitochondrial function) were used. Since the development of chronic diseases is almost always related to specific organs, the biological relevance of non-organ specific effects to predict the risk of developing of persistent damage is questionable.

Table 8.1 summarises the currently available models for long-term testing in relation to three of the most important targets for toxicity and/or sites of detoxification (the liver, kidney and central nervous system), indicating the advantages and disadvantages of each approach (Minuth *et al.*, 1992; Hopfer *et al.*, 1996; Schmuck & Schlüter, 1996; Houy, 1998; Hanley *et al.*, 1999; Pfaller *et al.*, 1999; Scheers *et al.*, 2001). Liver toxicity is considered in Section 7.4.

8.3 The ECVAM workshop on *in vitro* methods for long-term toxicity testing

An ECVAM workshop on novel advanced *in vitro* methods for long-term toxicity testing has been held, and the report and recommendations have recently been published (Pfaller *et al.*, 2001). The main conclusions regarding the development of *in vitro* repeat-dose toxicity tests are summarised in this section. The main recommendations of the workshop participants are included in the recommendations given below (Section 8.4.2).

1. One of the problems with *in vitro* toxicity testing is the lack of a precise definition of "long-term", as compared with *in vivo* approaches. "Long-term" in the *in vitro* context should be defined as meaning a minimum of 5 days, but the development of systems usable over several months is also necessary and feasible.
2. Relatively stable cultures are a prerequisite for long-term *in vitro* toxicity studies. Currently, systems such as brain-tissue slices, static cell cultures, transformed or immortalised cell lines, and genetically-engineered cell lines can meet this criterion.
3. The long-term maintenance of *in vitro* systems will require the use of new cell culture methods, which allow for the repeated or continuous administration of test compounds.
4. Due to the nature of long-term toxicity testing and the need to maintain *in vitro* systems with physiological characteristics similar to those in the *in vivo* situtation, it should be recognised that *in vitro* systems for chronic toxicity will be more complicated than *in vitro* systems for acute toxicity. This should not preclude the development of long-term *in vitro* systems, and their development should not be focused on their ability to permit high-throughput screening.
5. Attention needs to be paid to the definition, control and monitoring of cell cultures and culture conditions, taking into account organ-specific requirements.
6. Although IC50 determinations are suitable for acute toxicity, they are not necessarily relevant for long-term studies. It can be assumed that compounds that

are toxic over a short period will also be toxic over a long period. Therefore, long-term toxicity testing should be based on the determination of acute NOELs.

7. For the development of relevant and reliable *in vitro* procedures for long-term testing, it will be necessary to identify and use good reference compounds for specific kinds of tissues and test items.

8. Measurements obtained with *in vitro* models should ideally be validated by comparison with human data. Since it is not usually possible to measure effects *in vivo* (at least not non-invasively, and not for subtle changes), methods for detecting effects in organs should be based on non-invasive imaging techniques, or on measurements in surrogate tissues (such as blood, urine or faeces).

8.4 Repeat-dose toxicity: summary, conclusions and recommendations

A wide range of endpoints are investigated in *in vivo* chronic toxicity studies, so an integrated approach to chronic toxicity testing, based on the use of alternative methods with complementary endpoints, will need to be developed if the current reliance on chronic animal tests is to be reduced. At present, a variety of *in vitro* systems, derived mainly from the liver, kidney and brain, are being developed. Considerable investment at the research level is needed to maintain progress in this area.

8.4.1 Long-term prospects

The readiness of *in vitro* models for long-term effects to undergo prevalidation and validation will depend on progress made at the research and test development levels, for which recommendations are given below.

8.4.2 Recommendations for research and development

It is recommended that:

1. Initial efforts should focus on identifying compounds with well-established effects, for use as reference standards, and on tackling specific problems relevant to particular kinds of tissues and test items (ideally based on a knowledge of effects in humans), rather than on trying to replace chronic testing *in vivo* as a whole.

2. Attempts should be made to identify relevant biomarkers of exposure and effect. The endpoints selected should cover general cytotoxic mechanisms and cell-type-specific mechanisms of toxicity.

3. Wherever possible, human-based *in vitro* cell systems should be used.

4. Test development should be focused on the use of:

a) non-invasive imaging techniques for detecting long-term effects *in vivo*;

b) simplifed test parameters, such as electrophysiological applications for neurotoxicological long-term models, and the use of specific dyes or other markers, which permit changes in cell function to be monitored;

c) perfusion culture systems, with an emphasis on miniaturisation and practicability, to provide effective, technically simple and sensitive systems for assessing the effects of test compounds with biologically relevant endpoints;

d) metabolically competent, genetically engineered cell lines, grown on microporous supports and continuously perfused with conventional culture medium, since these

systems represent promising models for evaluating the effects of continuous low doses and long-term exposure in the liver, kidney and neuronal tissue;

e) methods employing human-based hepatic, renal and neuronal cell lines expressing a wide range of drug and xenobiotic metabolising enzymes and transport molecules, as an alternative to primary cultures;

f) long-term culture methods (several weeks to months) for hepatic and renal epithelial and endothelial cells, as well as for neurons and glial cells;

g) co-culture systems: i) neurons with glia cells; ii) hepatocytes with monocytic, Ito cells and/or endothelial cells; iii) renal epithelial cells with renal microvascular endothelial cells; and iv) renal glomerular mesangial cells with glomerular endothelial and/or glomerular epithelial cells;

h) liquid-gas phase organotypic models for culturing epithelial lung cells; and

i) systems for predicting toxicotolerance.

5. It is also recommended that the following activities are carried out:

a) an assessment of the immunological basis of target organ toxicity.

b) an assessment of the usefuless and application of genomics and protemics in repeat-dose toxicity testing.

8.5 Current status of nephrotoxicity testing

The kidney is a highly complex organ, composed of many different types of cells, including epithelial, endothelial and interstitial cells. Its main function is the regulation of both the volume and composition of the extracellular fluid by regulating the excretion of solutes and water. The kidneys receive about 20% of the resting cardiac output, they are metabolically active organs, and they reabsorb, metabolise and secrete a large number of compounds. They are therefore highly susceptible to the effects of chemical toxicants. Toxic compounds affect mainly the cells lining the vasculature and the epithelial cells of the proximal tubules of the nephrons, which represent 60% of the total renal cell mass. Tubular damage and loss of epithelial integrity result in a decrease in the re-absorptive capacity of the kidney (Hawksworth *et al.*, 1995; Morin *et al.*, 1997).

The use of cell culture techniques to study nephrotoxicity *in vitro* has recently gained importance, due to improvements in methods used for growing homogeneous cultures of renal cells. Several *in vitro* methods are being investigated as models for nephrotoxicity (e.g. Bach *et al.*, 1996; Pfaller & Gstraunthaler, 1998). Table 8.2 provides a summary of the methods available, with their advantages and disadvantages.

Primary cultures of glomerular mesangial and epithelial cells from various sites along the nephron and permanent renal epithelial cell lines can be employed. Primary cultures are difficult to obtain, due to the large number of different cell types present in the kidney. These cells tend to dedifferentiate within hours, sometimes because of a lack of suitably specific culture media. In addition, there is not enough knowledge with respect to markers that can be used to monitor the maintenance of differentiation *in situ*. Despite all these problems, primary cultures maintain characteristics closer to the *in vivo* situation than do cell lines. Nevertheless, renal epithelial cell lines have proven to be a very powerful tool for use in studying important aspects of nephrotoxicity *in vitro*. They retain a number of differentiated functions of their *in*

vivo ancestor cells, they have unlimited life-span, and they do not require time-consuming isolation procedures (see Gstraunthaler, 1988).

A number of *in vitro* endpoints are being investigated. For example, the measurement of barrier function is a new criterion for assessing cell injury in epithelial monolayers that separate fluid compartments (see Section 7.2 and Steinmassl *et al.*, 1995; Pfaller & Troppmair, 2000). An ECVAM study on the use of trans-epithelial resistance and inulin permeability as nephrotoxicity endpoints in LLC-PK1 and MDCK cells is in progress. The aim of the study is to establish whether transepithelial resistance and trans-epithelial marker permeability can be used as general predictors of nephrotoxicity in intact renal epithelia *in vitro,* by investigating the actions of a range of chemicals which have been identified as nephrotoxic *in vivo*.

8.6 Nephrotoxicity: summary, conclusions and recommendations

Renal trans-epithelial resistance and paracellular permeability are reproducible endpoints for assessing barrier function.

8.6.1 Short- and medium-term prospects

1. Investigations on the use of tissue renal slices as a model for studying the mechanisms of nephrotoxicity.
2. The immortalisation of primary proximal tubular cell cultures, and further characterisation of the new cell lines, to establish whether they can retain the characteristics of their *in vivo* precursor cells through several passages.
3. The further development of the use of the HK-2 (human proximal tubular epithelial) and LLC-PK1 cells (porcine proximal tubular epithelial) cell lines, grown under long-term conditions, and assessment of their usefulness for transport studies.
4. Investigations into the use of molecular biological techniques to re-express lost functions, e.g. specific transporters, enzymes or receptors.

8.6.2 Long-term prospects

1. The further development of the available renal co-culture systems (e.g MDCK cells with metabolically-competent cells), and the development of human renal systems, since epithelial and endothelial cells can be successfully grown in combination under static culture conditions and under continuous medium perfusion.
2. An assessment of the extent to which loss of barrier function can account for kidney damage.

8.7 Current status of neurotoxicity testing

Neurotoxicity is a major form of target organ toxicity that can result in lethality. The current OECD Test Guidelines 418 and 419 (OECD, 1995a &1995b) and Test Guideline 424 (OECD, 1997) for assessing the neurotoxic effects of chemicals are based on *in vivo* studies. There are a number of short-falls in the current whole-animal approach for evaluating the safety of chemicals, including escalating costs, slow throughput of compounds, and increasing animal usage, as a result of the increasing number of chemicals being developed and commercialised (Atterwill *et al.,* 1994;

Costa, 1998). In addition, the tests incorporated into the current guidelines do not always generate the mechanistic data required for a scientifically-based human risk assessment (ECVAM, 1994). Therefore, the need to develop mechanistically relevant alternatives to conventional animal toxicity testing is widely recognised (Abdulla *et al.*, 1995; Costa, 1998). The best way forward will be to design and assess the use of tiered strategies based on batteries of tests that encompass the most important neurotoxic endpoints. To date, no single *in vitro* method for neurotoxicity has been validated, and no battery or testing strategy has been independently evaluated. A major challenge is to identify a short-list of appropriate neurospecific endpoints to be incorporated into a testing strategy.

8.8 Testing strategies for *in vitro* neurotoxicity

A tiered approach is widely considered to give an optimum balance between the accurate detection of possible neurotoxins and ease, speed and the efficient use of resources (Atterwill *et al.,* 1994). Although consensus has not been reached on the specific composition and design of an optimal testing strategy, a commonly recommended approach is to employ at least two tiers of tests. Typically, a first tier should enable specific neurotoxicants to be distinguished from general cytotoxicants, by using combinations of non-specific and neurospecific endpoints in neural and non-neural cell lines. A second tier should enable different classes of neurotoxicants to be distinguished, and should address more-specific mechanistic questions. The second tier is therefore likely to involve the use of specific neural endpoints, such as the activities of acetylcholinesterase (AChE) and other enzymes involved in neural function. A number of testing strategies have been proposed (Veronesi, 1992; Walum *et al.*, 1992; Abdulla & Campbell, 1993; Atterwill *et al.*, 1994; Williams *et al.*, 1994; Pentreath & Atterwill, 1996; Costa, 1998; Balls & Walum, 1999). Many of these schemes are complex, but they make useful suggestions for future development.

To provide a means of linking the external dose of a chemical to the target organ concentrations that produce toxicity, a generic testing strategy for neurotoxicity has been based on the integrated use of biokinetic and biodynamic parameters (Blaauboer *et al.,* 1999). The proposed scheme is based on a measurement of basal cytotoxicity, combined with the use of biokinetic models to predict the dose of a chemical which could result in sufficiently high concentrations to cause non-specific effects, for example, in the central nervous system (CNS) or peripheral nervous system (PNS). In addition, it is suggested that the chemical should be tested by using a battery of neuron-specific *in vitro* tests. If a high ratio of general cytotoxicity to specific cytotoxicity is observed in these *in vitro* systems, and if biokinetic considerations indicate an accumulation of the chemical at the target site in the nervous system, the chemical of concern could be regarded as an acute or chronic neurotoxicant.

8.8.1 Tests for general neurotoxicity

Tests used for screening purposes should be simple and designed to detect a broad array of chemical insults to the nervous system. They should be able to detect compounds that may be toxic to either the adult or to the developing CNS and PNS.

Although many different approaches have been reported (e.g. Atterwill *et al.*, 1994; Fielder *et al.*, 1997; Pentreath, 1999), the measurement of cytotoxic endpoints in human cell lines, such as inhibition of NRU and MTT reduction, is generally

proposed. Some preliminary feasibility studies have already been carried out. For example, Williams *et al.* (1994, 1996) studied a total of 43 chemicals, by using a battery of neuroblastomas, primary neural and astrocyte cultures, and fibroblasts as non-neural cells. The endpoints analysed were MTT reduction, NRU and LDH release. The results indicated that a battery composed of human IMR-32 neuroblastoma cells, fibroblast cultures and primary astrocytes, did not enable discrimination between neurotoxicants, gliotoxicants, and general cytotoxicants. However, the use of human IMR-32 neuroblastoma cells, with NRU as an endpoint, emerged as a suitable model for inclusion in a more-comprehensive first-tier screen for neurotoxicants.

Weiss & Sawyer (1993) tested 50 MEIC reference chemicals in primary cultures of chick embryo forebrain neurons, by using the MTT and NRU assays. The NRU assay appeared again to be more sensitive to chemical toxicity; however, both assays were equally predictive when compared with *in vivo* toxicity data obtained from the Registry of Toxic Effects of Chemical Substances (RTECS).

Finally, Xie & Harvey (1993), tested the first nine MEIC chemicals in the neuroblastoma NG108-15 cell line, by using the MTT assay to measure cell viability, and by monitoring changes in cell resting membrane potential (RMP) as a neurospecific marker. The authors concluded that the MTT assay may be a better approach than RMP measurement for detecting neurotoxic effects *in vitro*, because even simple electrophysiological techniques are not suitable for rapid screening, and because there was a strong correlation between the rank order of potencies of the chemicals in the MTT and RMP assays (Balls & Walum, 1999).

8.8.2 Tests for mechanism-specific neurotoxicity

Mechanism-specific neurotoxicity tests are useful when general *in vitro* screening tests are not capable of detecting the effects of specific chemical classes (Abdulla *et al.*, 1995). In a series of papers by Veronesi and co-workers (for review, see Balls & Walum, 1999), the use of neuroblastoma cells as effective test models for organophosphate neuropathy was demonstrated. It was found that the human SH-SY5Y cell line could identify active esterase inhibitors among the organophosphates tested, and could also distingush between organophosphates that caused delayed neuropathy and those that did not. The authors therefore concluded that there was a distinct possibility that a human cell line could be used for the screening of organophosphates, and that neuropathy target esterase (NTE) inhibition could be identified without using laboratory animals. In another study, Henschler *et al.* (1992) were able to establish an almost perfect correlation between *in vitro* and *in vivo* data for a series of organophosphorus compounds. The *in vivo* data were obtained with the standard *in vivo* hen test, whereas the *in vitro* data were obtained by observing and quantifying the development of neurite-like processes in neuroblastoma (N-18) and glioma (C6) cell lines.

Forsby *et al.* (1995) analysed specific neurotoxicity on synaptic events, by using Ca^{2+} homeostasis as an endpoint with human neuroblastoma cells (SH-SY5Y). In their recent review, Balls & Walum (1999) concluded that although perturbation of calcium homeostasis may be a common denominator in many neurotoxic conditions, the determination of intracellular calcium concentrations in resting and activated cells is not sufficient for identifying a broad range of neurotoxic compounds. *In vitro*

neurotoxicity tests for assessing synaptic activity should therefore be based on, for example, determinations of the activity of ligand-operated ion channels, adenylate cyclase-coupled receptors, and neurotransmitter transporters, and on measurements of transmitter release, oxidative stress and stress gene activation.

Among the various *in vitro* models proposed as second-tier tests for the assessment of neurotoxicity (Table 8.3), reaggregating brain cell cultures appear to be among the most promising models. These systems consist of three-dimensional, integrated populations of neurons and glial cells, derived from fetal rat telencephalons. The advantages of the three-dimensional model for neurotoxicity testing are: a) they are robust; b) they can be used in both acute and chronic exposure studies; c) they permit testing in both immature and differentiated cells; d) they reproduce most of the cell-cell interactions that occur *in vivo*, including morphogenetic steps such as proliferation, synaptogenesis and myelination; and e) they can be used for studying a wide range of specific neurotoxic endpoints, including both microglial and astroglial activation, which are early markers of neurotoxicity (Atterwill & Purcell, 1999; Monnet-Tschudi *et al.*, 1997). Reproducible data show that adverse effects observed in aggregating brain cell cultures occur at similar concentrations to those observed *in vivo* (Honegger & Schilter, 1992; Kucera *et al.*, 1993; Zurich *et al.*, 2000).

One drawback of re-aggregating brain cell cultures is the need to work with primary cells. However, it appears that logistical difficulties associated with the availability and distribution of reaggregating cultures can be overcome, since preliminary results on the possible cryopreservation of the aggregate cultures are very promising. Other limitations are the incomplete reconstitution of the *in vivo* cytoarchitecture, and the need for specialised equipment.

In the long-term, human embryonic cell lines are likely to be promising *in vitro* models for a wide range of purposes. Such systems are advantageous in that they avoid the difficulty of extrapolating between species and avoid the use of animals. A human embryonic cell line has been shown to be capable of differentiation, under controlled conditions, into the various classes of neural cells, neurons, astrocytes, oligodendrocytes and microglia (Svendsen *et al.*, 1998). Such systems may therefore reach a level of biological complexity comparable to aggregate cultures. However, they need to be further developed and assessed for their relevance to the *in vivo* situation.

Another model that holds promise for the detection of mechanistic pathways of neurotoxicity is the genetically-modified phaeochromocytoma 12 (PC12) cell line (Stingele *et al.*, 1999). Two classes of genetically-modified PC12 cell lines have been patented by ECVAM, which differ in terms of their sensitivity to P53-mediated and Bcl2-mediated cell death. Currently, the usefulness of these cells as a part of a testing strategy for neurotoxicity is being assessed by ECVAM.

8.9 Neurotoxicity: summary, conclusions and recommendations

A large number of *in vitro* systems are being developed to investigate the wide variety of endpoints associated with neurotoxicity. This is an area where a reduction in animal use is likely to result from the integrated use of alternative tests with complementary endpoints. An important question in relation to the chemicals policy is

the extent to which all mechanisms need to be assessed, since, from a strategic viewpoint, it is possible that an assessment of barrier function (see Section 7.2.5 on the BBB), combined with assessments of basal cytotoxicity (Chapter 2) and energy metabolism, might be sufficient to identify substances of concern.

8.9.1 Short-term prospects

1. The prevalidation of the complementary use of re-aggregating brain cell cultures and the human neuroblastoma SHSY5Y cell line.
2. The prevalidation of the complementary use of primary glial and neuronal cell cultures for assessing neurotoxicity.
3. An evaluation of the usefulness of genetically-modified PC12 cell lines.

8.9.2 Recommendations for research and development

1. The development and evaluation of a tiered testing strategy for neurotoxicity, taking into account the results obtained in previous studies (Walum *et al.,* 1992; Atterwill *et al.*, 1994; Williams *et al.*, 1994; Williams *et al.*, 1996; Pentreath & Atterwill, 1996; Costa, 1998; Balls & Walum, 1999).
2. The further development of *in vitro* models for evaluating the mechanisms of neurotoxicity, such as genetically engineered cell lines, and re-aggregating cultures of human embryonic stem cell lines.
3. The evaluation of genomics, proteomics and new electrophysiological and biochemical profiling methods for neurotoxicity testing

Table 8.1. An overview of models for chronic toxicity testing.

Models	Advantages	Limitations
Hepatotoxicity Isolated perfused liver	The *in vitro* model closest to the *in vivo* situation	Short life-time (2-3 hours) Complicated and demanding set-up
Hepatotoxicity Liver slices	Retains the *in vivo* tissue organisation	Short life-time (7 days)
Hepatotoxicity Isolated hepatocytes	The most frequently used *in vitro* model for long-term hepatotoxicity testing Can produce a metabolite profile of a drug very similar to that found *in vivo*	Short life-time (24 hours) Loss of many liver specific functions Availability of human cells very limited Non-human models are not always predictive of human *in vivo* situation
Hepatotoxicity Collagen sandwich cultures	Structural and functional integrity retained for up to 15 days Normal cell shape is kept; Intact structure, bile canaliculi are preserved ALAT and ASAT enzyme releases can be studied Successfully used in mimicking the chronic treatment *in vivo* situation	Loss of several differentiated functions over time
Hepatotoxicity Ito cells and Ito cells/hepatocytes co-culture	Useful tools for studying liver fibrogenesis; Life-time 96 hours	Not a well-established model

Hepatotoxicity Genetically-engineered cells expressing single human or animal P450 enzymes	Used as a tool to assess the involvement of certain enzymes in metabolism, metabolite formation, and metabolism-dependent toxicity.	Well-characterised, well-documented and easy to use
Hepatotoxicity Cell lines derived from human hepatoma	Expression of CYP1A1	Low expression of other CYPs
Hepatotoxicity HepG2 grown under continuous medium supply	Only used in a few pilot studies	Laborious systems for routine use.
Nephrotoxicity Renal epithelial cell lines grown under continuous medium supply	Phenotypes with oxidative energy metabolism are available for tubular proximal cells Morphology very close to the *in vivo* parent cell type Culture periods of up to 6 weeks easily possible; Human derived proximal tubular cell lines similar in function to parent cells are available Many functional parameters can be monitored and a number of endpoints can easily be assessed	Laborious systems for routine use.
Neurotoxicity Primary neuronal cell cultures (rat)	Well characterised Identification of neurodegenerative compounds Long-term exposure possible (7-14 days) Used in industry for screening pharmaceuticals and agrochemicals Possibility to study oxidative stress and excitotoxicity	
Neurotoxicity Permanent neuronal cell lines	Useful for detection of delayed neurotoxicity caused by organophosphates	Only one model available

Neurotoxicity Astrocytes	Culture system well developed Require long-term culture for differentiation Possibility to detect reactive astrocytes (cytokine production) Specific surface markers can be detected by cell imaging and FACS analysis after a toxic insult.	
Neurotoxicity Oligodendrocytes	Need long-term culture for maturation and/or testing of toxic compounds *In vitro* models for studying dysmyelination and demyelination are established Specific surface markers can be detected by cell imaging and FACS analysis after a toxic insult.	
Neurotoxicity Microglia	Rapid response to neuronal injury Cytokine production, morphology, phagocytosis and proliferation are useful endpoints Cultures require 1-2 weeks for the production of well- characterised microglia Cell culture models and endpoints measurements are successfully developed Primary glial cell cultures are widely used in academia and in industry	
Neurotoxicity Brain slices from hippocampus	Useful for detecting excitotoxic and/or convulsive properties of drugs (hardly detectable in rodents) Well-accepted for studying learning and memory deficits (difficult to detect in rodents) Used in agrochemical testing Long-term culture conditions are required	

Table 8.2.   The current status of models for nephrotoxicity testing.

Models	Advantages	Limitations
Higher order systems Rat or rabbit isolated perfused kidney	Tubulovascular integrity is preserved	Animal model Not useful for routine examination Renal function is maintained only for a short period
Higher order systems Isolated perfused nephrons or nephron segments	Allows functional characterization (transport, electrical properties) of nephron segments Useful to localise enzyme systems, metabolic pathways, and receptor distribution	Animal model Not suitable for routine *in vitro* studies
In Vitro Renal tissue slices	Allows to study of transport and toxicity Good results with human kidney slices Interesting tools to study nephrotoxicity, their mechanisms and prevention	Contain heterogeneous cell population There is a complex system Cells and exposed surfaces damaged from cutting Short life span.
In Vitro Isolated glomeruli, tubular fragments, and renal cells	Assessment of acute effects of chemicals Retention of *in vivo* architecture of the epithelium Useful to study unspecific but sensitive endpoints of toxicity testing Study of nephrotoxins in metabolism and transport	Life-span limited to a few hours
In vitro Primary cultures of glomerular and epithelial cells	Close to *in vivo* situation Use of microporous supports to grow the cells improves differentiated state during prolonged times	Limited life-span Isolation procedure is difficult Rat proximal tubules are very difficult to grow
In vitro permanent cell lines LLC-PK1 (proximal	Unlimited lifespan Possibility of long-term exposure under controlled conditions	Dedifferentiation occurs in culture Possible transdifferentiation occurs

nephron) OK (proximal nephron) JTC-12 (proximal nephron) MDCK (collecting duct) A6 (distal tubule/collecting duct)	Rapid isolation is possible Appropriate culture conditions available for continuous culture medium perfusion, well-defined extracellular matrix, and new technologies can be used to induce re-expression of lost functions Gluconeogenic strains of OK and LLC-PK1 cells have been obtained	

Table 8.3.   An overview of alternative test batteries proposed for neurotoxicity testing.

Strategy	Test systems	Endpoints
Balls & Walum (1999)	*First tier* Structure-activity relationships Pharmacokinetic models *Second tier* Human cell lines *Third tier* Human cell lines *Fourth tier* CNS and PNS models of animal origin, including: brain spheroids, dorsal root ganglia, explants, dissociated neuronal and glial primary cultures, and microglial/mast cells	 *Second tier* Cell viability tests (necrosis and apoptosis) *Third tier* Class-specific (e.g. organophosphate) toxicity Axonal degeneration Cell signalling alterations Neurosecretory function Markers of neuroimmunotoxicity *Fourth tier* Glial fibrillary acidic protein (GFAP) Myelin basic protein (MBP) Neurofilament protein (NFP) Markers of excitotoxicity Markers of oxidative stress Morphological and neurochemical markers of synaptic damage Markers of neuroimmunotoxicity

Atterwill *et al.* (1994)	*First tier* Primary culture enriched in neurons Primary culture enriched in astrocytes Primary culture of dorsal root ganglia (for PNS toxicity) Clonal cell lines of neural origin	*First tier* MTT reduction NRU Fluorescein diacetate hydrolysis Cell-specific markers such as GFAP, NFP Cell growth, proliferation, motility, adhesion and process formation
	Second tier Whole-brain reaggregate cultures Clonal cell lines of neural origin	*Second tier* GFAP NFP MBP Malondialdehyde assay Calcium levels
	Third tier New models	*Third tier* New endpoints
Williams *et al.* (1994)	*First tier* Human neuroblastoma (IMR-32 or SHSY-5Y) Pure rat astrocyte cultures or glioma cells Fibroblasts (i.e. non-neural cells)	*First tier* NRU MTT reduction Trypan blue exclusion Total protein content AChE activity NTE activity
	Second tier Rat embryo whole-brain reaggregate cultures	*Second tier* Diameter Total protein content

		LDH release Neuron-specific enolase GFAP Neurotransmitter markers
	Third tier Pure neurons Pure astrocytes Human cell model	*Third tier* Cytotoxicity 2-deoxyglucose uptake/release GFAP/NFP
Costa (1998)	Neuronal and glial cell lines One non-neuronal cell line Rat primary micromass or reaggregate cultures	*Cytotoxic endpoints* Cell death Membrane permeability Mitochondrial function Cell growth and division Energy regulation Synthesis of macromolecules *Neurospecific endpoints* Neurotransmitter synthesising enzymes (glutamic acid decarboxylate, dopamine hydroxylase, choline acetyltransferase, tyrosine hydroxylase) Acetylcholinesterase Neuron-specific enolase Glutamine synthetase Neuronal receptors Neurite extension

| Abdulla & Campbell (1993) | *First tier*
Cell lines (neuroblastomas/gliomas)
Primary monolayer cultures (e.g. chick DRG)
Rat or chick midbrain micromass cultures
Microglia
PC12 cells

Second tier
Reaggregate culture (e.g. chick embryo)
Retinal cultures
Organotypic culture/explant culture
Primary cultures of individual neural cell types | *First tier*
Membrane integrity (e.g LDH release)
Mitochondrial enzyme function
Differentiation effects
CR3 expression (microglia to macrophages)

Neurite outgrowth
Gene expression (e.g. c-fos/c-jun)

Second tier
Aggregation (cell-cell interactions, gap junction formation)
Growth
Differentiation
Neurochemistry
Neurotransmitter synthesis (e.g. acetylcholine)
Receptor expression (e.g. GABA receptors)
Electrophysiology on *ex vivo* insect neurons |
| Veronesi (1992) US EPA | neuroblastoma cell lines (human SY5Y and murine NIE 115)

MCL-5 and AHH-1 TK+/- lymphoblastoids
Neuroblastoma cells co-cultured with muscle cells | *Neurotoxic endpoints*
Acetylcholinesterase activity
NTE activity
Histochemistry
Receptor status
Neurite formation

Cytotoxic endpoints
Organelle viability
Intercellular communication
LDH release |

9.   GENOTOXICITY AND CARCINOGENICITY

9.1 Introduction

A chemical is considered to be mutagenic if it is capable of inducing heritable changes (mutations) in the genotype of a cell as a consequence of alterations to, or loss of, genes, chromosomes or parts of chromosomes. Genotoxicity is a broader term that refers to the ability to interact with DNA and/or the cellular apparatus that regulates the fidelity of the genome, such as the spindle apparatus and topoisomerase enzymes (IEH, 2002).

A chemical is regarded as carcinogenic if it induces cancer (or increases the incidence of cancer formation). The process of carcinogenesis is considered to result from the transition of normal cells into cancer cells via a sequence of stages associated with alterations in rates of cell growth and death, cell differentiation and the invasion of healthy tissues by cancer cells. Genotoxic carcinogens induce cancer by a direct interaction of the chemical (or an active metabolite) with DNA, whereas non-genotoxic carcinogens exert their carcinogenic effects through mechanisms other than genotoxicity.

9.2 Current approaches to genotoxicity testing

Genotoxicity testing is an important part of the hazard assessment of chemicals for regulatory purposes (Fielder *et al.*, 1997; DOH, 2000). It is undertaken for two main reasons (DOH, 1989; Combes, 1995): a) to detect chemicals that might cause genetic damage, including point mutations, in germ cells, and thus increase the burden of genetic disease in the human population; and b) to detect chemicals that might be carcinogenic (based on the assumption that mutagenesis is a key event in the process of carcinogenesis). A scheme for genotoxicity testing is presented in Figure 9.1.

9.2.1 The use of in vitro assays in regulatory genotoxicity testing

The standard regulatory approach to genotoxicity testing is to use a tier-testing scheme comprising at least two *in vitro* tests in the first level of the tier: a bacterial mutagenicity assay (OECD TG 471; usually the Ames test with *Salmonella*, but sometimes with *E. coli*) and a cytogenetics assay (OECD 473; usually a metaphase analysis either with human lymphocytes or with rodent cell lines) (ICH, 1997; Mitchell & Combes, 1996; Combes, 1999; Zeiger, 2001). Two different endpoints are required, since chemicals can induce gene mutations and/or chromosomal damage. It is important to realise that these initial tests are used as a means of trying to establish the potential of a chemical to elicit a mutagenic effect, and that it is customary, particularly with drugs and food additives (i.e. where there is a likelihood of extensive and prolonged human exposure), to demonstrate that the mutagenic potential detected *in vitro* is realised in an intact animal. In other words, and contrary to hierarchical testing schemes for other types of toxicity, a negative result *in vitro* is usually considered sufficient to indicate lack of mutagenicity, whereas a positive result is not considered sufficient to indicate that the chemical represents a mutagenic hazard.

Some regulatory agencies encourage the use of the mouse lymphoma assay instead of a cytogenetics test, since this assay detects both point mutations and chromosomal damage. However, it is a difficult assay to perform, and it is necessary to conduct colony-sizing accurately, in order to be able to assess the ability of a chemical to induce chromosomal damage (Combes *et al.*, 1995).

It is also possible to use genetically-engineered cell-lines, expressing one or more phase I and phase II enzymes, to detect a variety of genotoxic endpoints caused by chemicals requiring metabolism (Combes, 1992; Macé *et al.*, 1997; Wiebel *et al.*, 1997). These have the advantage over exogenous metabolising systems that metabolites are generated intracellularly, i.e. close to the target DNA.

9.2.2 The use of in vivo assays in regulatory genotoxicity testing

The usual approach in *in vivo* genotoxicity testing is to look for cytogenetic damage in the bone-marrow by using either the micronucleus assay or metaphase analysis. *In vivo* tests are used to determine whether the potential for genotoxicity detected *in vitro* is realised *in vivo* (i.e. to identify false positives). False positives *in vitro* can have a variety of causes, such as preferential metabolic activation and excessively high dose levels, without transport away from the target site.

Despite the problems inherent in interpreting the data produced (see below), a negative result is often taken to infer that the chemical is unlikely to be either a germ cell mutagen or a genotoxic carcinogen, although further assessment, possibly culminating in a rodent bioassay, is sometimes required. A positive result in the bone-marrow is interpreted as being indicative of genotoxicity to the germ cells (as there are no known germ cell mutagens that are not somatic cell mutagens), and also of carcinogenicity via a genotoxic mechanism. In this case, and depending on likely exposure levels in humans, a rodent bioassay may be required.

9.2.3 The detection of genotoxicity in different target tissues

The intra-peritoneal route of administration is often used in short-term genotoxicity assays, supposedly to maximise the amount of compound entering the body, and to minimise the possibility that a negative result in the bone-marrow could have been due to lack of transport to this target site. It is also possible to assess systemic toxicity to the bone-marrow by assessing the ratios of normochromatic and polychromatic erythrocytes with and without chemical exposure. A negative result in a bone-marrow micronucleus test, accompanied by evidence of distribution to the bone-marrow, is evidence of lack of genotoxicity. However, when there is no indication of transport to the bone-marrow, a negative result is impossible to interpret.

Other *in vivo* assays are sometimes required, especially if evidence from pharmacokinetic studies indicates that tissues other than the bone-marrow are targeted. Such assays include the detection of unscheduled DNA synthesis (UDS) in the liver, and the use of transgenic rodent systems, developed to detect mutagenicity simultaneously in a wide range of tissues. However, these adjunct tests are not used routinely, and have not been validated according to recognised international criteria (Mitchell & Combes, 1996).

In vivo genotoxicity tests are unsatisfactory for reasons other than the above, in that they are restricted to a few tissues, and also they are limited to detecting effects caused by a narrow range of mechanims, such as cytogenetic damage and DNA repair (except in the case of the relatively new transgenic rodent mutagenicity models, that still require formal validation).

9.2.4 The detection of aneupolidy

Changes in chromosomal number (polyploidy or aneuploidy) can be investigated both *in vitro* and *in vivo* by using metaphase analysis and chromosomal painting techniques, although these are labour-intensive methods. Numerical changes in chromosome number will lead to alterations in the genotype and are therefore considered to be genetic changes. However, aneuploidy can arise both as a result of genotoxic and non-genotoxic events (as defined in this chapter), since loss of chromosomes can be due to either direct effects on the chromosome to produce an acentric fragment, or to interference with the site of attachment of the chromosome on the spindle.

Some authorities have called for the development and use of specific tests for aneuploidy (DOH, 2000). In relation to this, there have been efforts to develop an *in vitro* micronucleus test, although technical difficulties with this assay have delayed its validation, which is still in progress (Kirsch-Volders *et al.*, 2000). The availability of an *in vitro* micronucleus test would be a considerable advance on the current situation, since scoring micronuclei is considerably easier than is scoring chromosomal damage, thereby making the time and cost of *in vitro* cytogenetics testing equivalent to the *in vivo* assay. Also, micronuclei can be formed by genotoxic and non-genotoxic events, with the possibility of separately quantifying the amount of each event (Stopper & Müller, 1997; Aardema & Kirsch-Volders, 2001).

9.3 Current approaches to carcinogenicity testing

9.3.1 The rodent bioassay

The traditional approach to carcinogenicity testing is the life-time rodent bioassay in rats and mice of both sexes, with full pathological analysis of all tissues (OECD TG 451). This test is time-consuming, labour-intensive and very costly, both financially and in terms of animal numbers and animal welfare. The rodent bioassay is used to detect complete carcinogens, as well as tumour promoters and co-carcinogens (agents that can increase the frequency of tumorigenicity by acting in combination with another chemical or subsequent to another chemical). Moreover, rat and mouse data do not correlate well, and extrapolating the information to humans is problematic (Combes, 1997; Gottmann *et al.*, 2001). A scheme for carcinogenicity testing, based on genotoxicity assays and other tests, is presented in Figure 9.2.

9.3.2 The detection of genotoxic and non-genotoxic carcinogens

Data from the US National Toxocology Program (NTP) have shown that, in general, genotoxic chemicals are likely to exhibit trans-species carcinogenicity, often in both sexes, at intermediate dose-levels, and not necessarily restricted to one target tissue (Ashby & Tennant, 1991). These chemicals are the so-called "genotoxic carcinogens". The carcinogenicity of the vast majority of these chemicals could have been predicted by undertaking genotoxicity testing, without the need for the bioassay. However, an

increasing number of chemicals, devoid of any genotoxicity, are proving to exhibit carcinogenicity, and this activity is often manifested in one species, one sex and even in one specific tissue, and usually only at high dose levels (Choy, 2001). These chemicals are the so-called "non-genotoxic carcinogens".

The carcinogenic activity of non-genotoxic carcinogens cannot be predicted by the standard battery of genotoxicity assays (Combes, 2000b). In contrast to genotoxic carcinogens, which act by damaging DNA, non-genotoxic carcinogens are known, or thought, to act initially by a number of different mechanisms, some of them involving binding with specific intracellular receptors. At present, there few *in vitro* methods for detecting non-genotoxic carcinogens, and none of these has been validated for regulatory usage. Therefore, chemicals that prove to be non-genotoxic *in vitro* and *in vivo* may still be subjected to a rodent bioassay, depending on their intended use, and if a high level of human exposure is anticipated.

In recent years, the relevance of non-genotoxic carcinogens to human hazard has been questioned, in view of their restricted activity profiles, especially since their effects are often detected exclusively in the mouse. As a consequence, there have been discussions by groups such as the International Conference on Harmonisation (ICH) (see D'Arcy & Harron, 1997) for mouse transgenic assays to be developed and validated as an alternative to the mouse bioassay (Tennant *et al.*, 1999). Some of these assays have been designed specifically to detect certain types of non-genotoxic carcinogens, while others are specific for chemicals with genotoxic activity. However, in a collaborative study organised by ILSI, no single transgenic mouse system or combination of systems have yet emerged that exhibit a sufficiently high sensitivity and specificity for rodent carcinogenicity to be considered relevant and reliable for regulatory use (Robertson *et al.*, 2001). In fact, the results from the ILSI study raise several new questions: a) there appears to be an overlap between the activities of genotoxic and non-genotoxic chemicals in some assays, prompting questions with regard to the different modes of action of these types of chemicals; b) the mechanisms by which enhanced tumorigenesis arise in some of the models in response to carcinogens are not fully understood; and c) more work needs to be undertaken before standardised and optimised protocols become available (van Zeller & Combes, 1999). It is also the case that any testing strategy involving the combined use of a rat bioassay with a mouse transgenic assay to detect carcinogens, still suffers from the need to undertake rodent to human extrapolation.

9.3.3 Modelling carcinogenesis in vitro

It is well established that the carcinogenic process can be modelled *in vitro* on the basis of morphological cell transformation (Combes *et al.*, 1999). Several cell transformation assays based on immortalised or non-immortalised rodent cell systems were developed a long time ago, and have been shown to be capable of detecting some well-known animal and human carcinogens, as well as tumour promoters. In fact, many of the presumed stages of carcinogenesis, as well as the role of oncogenes, have been established as a result of *in vitro* studies on cell transformation.

The main endpoints of cell transformation, focus formation and the acquisition of the ability to grow in soft agar, result from a loss of contact inhibition and loss of anchorage-dependence, respectively. The relationship between these cellular endpoints, which can be detected and quantified in cell culture, and *in vivo*

carcinogenesis and malignancy, can be demonstrated by malignant tumour formation after the transformed cells are injected into animals.

Modelling the later stages of carcinogenesis is increasingly difficult. Thus, *in vivo* experiments of cancer cell invasiveness and metastasis are limited in the detail they provide, so it essential to implement *in vitro* studies. Good examples are the three-dimensional *in vitro* invasiveness assays of Mareel *et al.* (1979), Smolle *et al.* (1992), Fusenig & Boukamp (1998), as well as the refinement of the latter technique by DeVaney *et al.* (1997). In addition, tumor neoangiogenesis can be studied in the allantoic membrane of chicken eggs.

9.3.4 Cell transformation assays

Recently, new protocols have been developed for two cell transformation systems, the Balb C/3T3 and syrian hamster embryo (SHE) cell assays, which have improved the reliability and predictivity of these assays for detecting rodent carcinogens (Combes *et al.*, 1999; Tsuchiya *et al.*, 1999). These developments have also resulted in an increase in the database for the assays, and have provided evidence that the techniques can be performed in different laboratories with a high level of reproducibility. This is especially so in the case of the SHE cell assay, which exhibited a high level of sensitivity for rodent carcinogens in comparison with several other alternative assays to the conventional rodent bioassay, including transgenic mouse strains, in the recent ILSI collaborative study. In this study, the SHE cell assay was highly predictive of the rodent bioassay, and it also detected human carcinogens, but was poor at distinguishing between rodent and human carcinogens (Mauthe *et al.*, 2001). It is noteworthy that 10/35 non-carcinogens were positive, though the mechanisms involved in generating these false positives are unknown at present. The four non-genotoxic carcinogens in the study were positive when exposure was conducted for seven days (Robertson *et al.*, 2002).

An analysis of a database comprising 48 chemicals tested in the reduced pH protocol assay showed 85% concordance between transformation and *in vivo* rodent carcinogenicity, with values for sensitivity and specificity of 87% and 83%, respectively (LeBoeuf *et al.*, 1996). Since these reviews were published, additional data have become available under the US NTP collaborative study and additional studies, for an analysis of data on a total of 75 chemicals tested at pH 6.7 to be undertaken. The results revealed an overall concordance of 83%, a sensitivity of 83% and a specificity of 82% (LeBoeuf *et al.*, 1999).

9.3.5 Screening for non-genotoxic carcinogens

The results of the ILSI study and other investigations have shown that some of the cell transformation assays are sensitive to a wide range of both genotoxic and non-genotoxic carcinogens, especially those that act via epigenetic effects and as tumour promoters and might affect cellular proliferation. Also, a two-stage cell transformation assay in Balb C/3T3 cells seems potentially useful for detecting tumour promoters (Tschuyi *et al.*, 1999). Morphological cell transformation has been shown to be due to point mutation, chromosomal damage, aneuploidy and other effects associated with cellular proliferation. However, due to the wide range of mechanisms by which non-genotoxic carcinogens might act, and especially as some of the effects are highly tissue-specific (e.g. thyroid carcinogenesis and peroxisome

proliferation), it is unlikely that a combination of this assay and one for aneuploidigenic agents would be sufficient to detect all types of non-genotoxic carcinogens.

Thus, to increase the spectrum of non-genotoxic carcinogens that can be detected *in vitro*, it will be necessary to develop a battery of assays that involve the detection of the principal endpoints by which these agents act. For example, there are several cell culture methods for investigating the tissue-specific induction of proliferation, such as an MCF-7 cell assay, that are being used to detect hormonally-active chemicals and the links with breast cancer (Combes, 2000a). In addition, several of the receptors for non-genotoxic hepatocarcinogenesis have been identified, which will facilitate the development of rapid ligand-binding assays (Combes, 2000b).

9.3.6 The regulatory status of cell transformation assays

There are several reasons why regulatory agencies have been reluctant to adopt cell transformation assays for routine testing. These reasons include: a) the perception that the assays are technically demanding; b) variability of results among different laboratories; c) uncertainties about data interpretation, and scepticism about the direct relationship between the cellular endpoint and tumour formation *in vivo*; d) a lack of information regarding the predictivity of some of the assays for rodent and human carcinogenesis; and e) the lack of formal validation of any of the cell transformation systems. It is mainly for these reasons that cell transformation has remained a tool for research on mechanisms of action of chemicals on a case-by-case basis, rather than becoming a regulatory approach for general application.

The problem regarding the demanding nature of the cell transformation assays needs to be addressed as a matter of urgency. Thus, although the low pH protocol for the SHE cell assay can be performed in different laboratories, it is considerably more complicated than the protocols required for other rodent cell transformation assays. Thus, it is suggested that all cell transformation assays should be developed further before being evaluated for their usefulness in predicting carcinogenicity.

Currently, few regulatory authorities accept cell transformation data for anything more than providing mechanistic information. For example, the US FDA indicates that cell transformation data can aid in compound selection and give useful mechanistic information, even in the case of negative results. Clearly, there is an important need for regulatory bodies to assess the current status of rodent cell transformation assays and their suitability for the detection and characterisation of carcinogens.

Human cells have been shown to undergo morphological transformation in tissue culture, and there is the possibility that reliable and sensitive human cell-based transformation assays could be developed in the not-too-distant future. Clearly, the eventual development of a human cell transformation assay would obviate the need for species extrapolation (Combes *et al.*, 1999).

9.4 SARs for predicting genotoxicity and carcinogenicity

Considerable advances are being made in predicting genotoxicity and carcinogenicity by developing and analysing databases and by applying computer-based approaches.

Diverse structural alerting features of genotoxic compounds have been recognised, particularly from the results of the US NTP collaborative study on short-term tests for carcinogenicity, and these have been used in predicting carcinogenicity, especially via genotoxic mechanisms (Ashby & Paton, 1993). By using fragment analysis, some 18 different structural alerts (toxicophores) have been identified as being present on chemicals that had been shown to be rodent carcinogens in two-year rodent life-time bioassays, as well as possessing genotoxicity in one or more short-term genotoxicity assays. However, fragment analysis approaches pay insufficient attention to the effect that one substituent group may exert on another, and there will be an increasing need for more-sophisticated methods, such as QSAR methods and three-dimensional modelling, coupled with knowledge-based expert systems (Barratt, 2000).

9.4.1 Expert systems

The computer-based expert systems for predicting the genotoxicity and/or carcinogenicity of chemicals are based on a variety of techniques, but essentially are divisible into automated rule-induction systems (ARI) and knowledge-based (KB) systems (Combes & Judson, 1995; Dearden *et al.*, 1997). These systems differ fundamentally in the way they operate, in that ARI systems predict by learning from and discovering patterns in existing data, whereas KB systems predict by reasoning on the basis of existing human knowledge. ARI systems make quantitative predictions (e.g. by providing a probability value that carcinogenicity will be induced by a molecule).

Three widely-used ARI systems are Topkat (Toxicity Prediction by Computer-Assisted Technology), CASE (Computer Automated Structure Evaluation) and COMPACT (Computerised Optimised Parametric Analysis of Chemical Toxicity). Examples of KB systems are HAZARDEXPERT and DEREK (Deductive Estimation of Risk from Existing Knowledge), both of which are used for predicting a wide variety of toxicity endpoints, and ONCOLOGIC, which is based on a decision-tree approach. There are also some KB systems for predicting metabolism, one example of which is a new program called METEOR.

ARI systems analyse information on structures entered into the computer for associations between inactive and active molecules. Molecules are fragmented into all possible atom pairs and other associations. Then, pattern recognition and cluster analysis techniques are used, together with statistical analyses, to compare the frequencies of occurrence of specific structural features in sets of active and inactive molecules, for a particular biological endpoint, by using a training set of chemicals. In this way, the most important features determining or modifying activity are identified. Such features are then searched for in novel molecules, which are entered into the system after it has been trained, and these are then processed for the presence of biophores and biophobes. Some ARI systems also use QSAR and molecular modelling of three-dimensional structure. There have been numerous papers showing how programs such as TOPKAT and CASE can predict the genotoxic and carcinogenic activities of chemicals (see, for example, Klopman & Rosenkranz, 1994).

COMPACT differs from these programs in that it analyses the ability of a molecule to fit into the active site of the of CYP 1A1 (and some other CYP isozymes), by modelling molecular shape (planarity or area/depth) and chemical reactivity (covalent

bond formation). The use of COMPACT is, of course, limited to molecules which are activated by these CYP enzymes.

DEREK is now one of the most widely used KB systems, and has an extensive rule-base for genotoxicity, especially for mutagenicity, initially based on the so-called Ashby alerts (see also Section 6.2.2). However, this rulebase has been added to extensively and refined, especially via feedback from the DEREK-users group (Ridings *et al.*, 1996). The rulebase for carcinogenicity has also been continually refined, particularly by the recent re-writing of the FDA rulebase that applies structural alerts to all epoxide substructures. Some mutagenicity and carcinogenicity rules have also been written for chemicals found in foods (Long & Combes, 1995).

9.4.2   The validation of expert systems for genotoxicity and carcinogenicity

There have been few studies in which expert systems have been compared for their ability to correctly predict the same biological activity, except in the case of rodent carcinogenicity, by using the NTP database (Parry, 1994; Lewis & Langley, 1996; Lewis *et al.*, 1996; Benigni, 2000). In these studies, several of the systems discussed above showed overall accuracies in correctly identifying rodent carcinogens which varied from 60-90 %, depending on the system and the database. Optimal levels of performance were obtained by using combinations of the systems. It can be concluded from these kinds of studies that expert systems should be used as screens in conjunction with each other and with *in vitro* tests. However, it is most important that expert systems should be validated according to the principles and procedures applied to other test methods. Unfortunately, this has not yet happened and, in some situations, systems such as COMPACT and CASE are being used only by one research group.

Expert systems are developed by using training sets of chemicals. It is important that such training sets are based on chemicals of varying structure and with a range of biological activities. Training sets should also contain chemicals acting via as wide as possible a range of mechanisms in causing the toxic effects of interest. It is also crucial that systems are developed that can correctly predict not only the activities of the chemicals in the training set, but also those of different chemicals, which are structurally-related to those in the training set. This is so that the utility of the expert system for predicting the activities of novel chemicals is as comprehensive as possible.

9.4.3   The limitations of expert systems

Expert systems have several important limitations. Firstly, their development depends crucially on the availability of accurate and relevant biological data of high quality on individual chemical entities with well-defined structures. Unfortunately, no test sample is completely pure, and it is very important that information used to construct the rules for expert systems is derived from studies on test samples of high purity. Thus, strict criteria should be applied to the process of data acceptance for rule development, and ideally, the effects of metabolism should be known, or should be predictable.

9.4.4 Structure-activity relationships for non-genotoxic carcinogens

The mechanisms involved in non-genotoxic carcinogenesis are less well characterised than are those for genotoxic carcinogenesis. As a consequence, there is considerably less information available for generating useful rules for predicting this type of activity. This issue was discussed by Combes (2000a), who also gave examples of the identification of some structural features of molecules that might be associated with non-genotoxic carcinogenesis, including information on some hormonally-active chemicals (Combes, 2000b).

9.5 Tiered testing strategies for genotoxicity and carcinogenicity

The principal problem in the development of a strategy for the mutagenicity and carcinogenicity testing of existing and new chemical entities revolves around two main questions: a) whether it is necessary to use *in vivo* genotoxicity assays; and b) whether it is necessary to test chemicals for carcinogenicity in the rodent bioassay or an alternative (equivalent) assay in animals.

It can be argued that, if a chemical is genotoxic *in vitro*, it should be regarded as a genotoxin with the potential to induce germ-line mutations and carcinogenicity, irrespective of its intended usage. The justification for using *in vivo* tests is that the chemical could prove to be negative *in vivo*, due to detoxification (a process less likely to occur *in vitro*), transport away from the site of action, and exposure to more realistic, lower dose levels. However, *in vivo* genotoxicity assays are unsatisfactory for other reasons (see section 9.2.2). Moreover, chemicals that are intrinsically genotoxic (without the need for exogenous metabolic activation), could exert effects in tissues after administration at initial sites of contact, before reaching the liver. In addition, chemicals requiring activation *in vitro* could be converted to reactive forms by extra-hepatic metabolism in such tissues (Combes, 1992).

The above arguments support the view that *in vivo* genotoxicity assays provide data of dubious relevance, and that the careful use and interpretation of data from a battery of *in vitro* tests might be a better approach. However, it is the usual experience in industry that a significant proportion of chemicals exhibit genotoxicity *in vitro,* but prove inactive *in vivo,* after thorough testing. In other words, potentially useful and important chemicals could often be discarded unnecessarily, if *in vitro* test data only were used for regulatory purposes. This conclusion is based on the assumption that the relative insensitivity of *in vivo* assays reflects a lack of realisation of genotoxic potential in the whole animal. However, the scientific basis for the insensitivity of *in vivo* genotoxicity assays is not known.

It is therefore suggested that the need for further *in vivo* genotoxicity testing should be justified on a case-by-case basis, in relation to the stage of screening being undertaken, the justification for undertaking *in vivo* assays in relation to known target organ distribution, and the anticipated level of human exposure to the chemical.

A chemical that is negative *in vitro*, in at least a point mutation assay and a cytogenetics assay, should be tested in a further *in vitro* assay, which should include aneuploidy as an endpoint, since this is known to be an important pathway for the effects of several chemicals that are uniquely genotoxic *in vivo* (Parry & Parry, 1995; Parry *et al.*, 1995; Anon, 1997; DOH, 2000). A positive result in such an assay should

be considered as indicative of carcinogenicity via a non-genotoxic mechanism, and other evidence for this should be sought from traditional chronic and sub-chronic study data (see below), and from the use of other approaches (see below). Also, an ability to induce aneuploidy, coupled with evidence for transport to the germ cells, indicates the possibility of important germ cell effects, in which case the chemical should either be banned or subjected to marketing restrictions. It should also be a general rule that chemicals that are unequivocally genotoxic (after testing *in vitro* and *in vivo,* as appropriate) should not be subjected to carcinogenicity bioassays. Tiered approaches for the assessment of genotoxicity and carcinogencity are outlined in Figures 9.1 and 9.2, respectively.

9.6 Issues requiring further consideration

The standard battery of *in vitro* and *in vivo* genotoxicity assays should detect all genotoxic carcinogens, but they will fail to detect most non-genotoxic carcinogens. One way in which the latter class of carcinogen might be detected is by the use of a cell transformation assay. It is, however, unrealistic at the present time to expect the regulatory authorities to endorse the use of transformation for regulatory purposes without the formal validation of optimised test protocols. It is therefore a matter of priority that the newly-developed protocols for the Balb C/3T3 and SHE cell transformation assays should be subjected to formal validation.

At the same time, there needs to be an ECVAM workshop to discuss the whole topic of non-genotoxic carcinogenesis, and to include detailed discussions on: a) the known modes of action of non-genotoxic carcinogens; b) the available and required non-animal testing strategies for detecting and characterising them; and, most importantly, c) their relevance to human hazard. The last-named discussion should focus on the evidence supporting the human carcinogenicity of non-genotoxic carcinogens, and how specific effects in rodents can be extrapolated to human hazard. Several non-animal methods are available, ranging from (Q)SAR studies and *in vitro* proliferation assays to aneuploidy assays. A further perceived need for using the rodent bioassay or an equivalent assay for detecting non-genotoxic carcinogens is the possibility that these chemicals exhibit a threshold dose, in contrast to genotoxic carcinogens (Fan & Howd, 2001). The data from bioassays are therefore used for regulatory purposes, to set exposure limits below which carcinogenic risk is considered to be much reduced. Most regulators consider that dose-response data from *in vitro* assays cannot be used to establish threshold doses, and use this argument to justify the use of *in vivo* experiments for this purpose. However, the current approach could be replaced in the following ways: a) by evaluating whether certain classes of non-genotoxic carcinogens, such as peroxisome proliferators, are relevant to human hazard; and b) by the combined use of information from toxicokinetic and chronic toxicity studies, in which target organ doses can be established (Blaauboer *et al.*, 1999) in relation to early endpoints of non-genotoxic carcinogenesis, such as cell proliferation (Butterworth *et al.*, 1995).

Some regulatory agencies have, in fact, explored the possibility of using dose-response data from certain *in vitro* assays (e.g. *in vitro* aneuploidy) for risk-assessment purposes, and this issue would also need to be explored in the proposed workshop.

There are two further issues with regard to genotoxicity and carcinogenicity testing: a) the focus by industry and regulatory bodies on predicting rodent carcinogenicity; and b) the unrealistic expectations being raised regarding the advantages of approaches based on genomic and proteomics.

In the first case, discussions with toxicologists in industry have confirmed the view that there is a disproportionate emphasis on finding alternatives to *rodent* bioassays, since this is the traditional route in regulatory toxicology. This approach is erroneous, since, ultimately, we should be interested in human carcinogens, and the relevance of rodent bioassay data to humans is questionable. This situation prompts the use of rodent *in vivo* data for validation studies on alternative methods, with insufficient use of the available human and mechanistic data.

In the second case, it is becoming increasingly clear from the outcome of several recent conferences and workshops, as well as from the literature, that expectations of genomics and proteomics have far exceeded the ability of these approaches to provide useful and reliable data that can be interpreted in terms of standard toxicity endpoints (Fielden & Zacharewski, 2001; Fry & George, 2001). This has largely been due to the way in which the sensitive molecular techniques have been applied. As a result, there have been projections that it will be at least 10 years before any useful information can be derived from genomics and proteomics. This needs to be recognised and acknowledged by the regulatory authorities and funding bodies, so that efforts to develop more-conventional *in vitro* toxicity methods and tests are not reduced in the expectation that genomics and proteomics will provide all the answers in the near future.

9.7 Genotoxicity & carcinogenicity: summary, conclusions and recommendations

QSAR, expert system and *in vitro* approaches could be used immediately to prioritise chemicals for further testing on the basis of their potential genotoxicity and carcinogenicity. For the definitive assessment of a genotoxicity, negative results in a bacterial assay for gene mutation and in a mammalian assay for clastogenicity/aneugenicity would normally be regarded as sufficient evidence for lack of genotoxic potential

For carcinogenicity testing, the rodent bioassay is not suitable for the testing of large numbers of chemicals, not only because of the considerable time and cost involved, but also because of the large numbers of animals required. Therefore, it is important that alternative methods for the definitive assessment of carcinogencity are further developed and then subjected to validation.

9.7.1 Short-term prospects

1. A review of the current validation and regulatory status of aneuploidy and micronucleus assays in mammalian cells, to include an assessment of the range of aneugens detected by both methods, with a view to eliminating redundancy.
2. A review of the current validation and regulatory status of rodent cell transformation assays (especially the SHE cell and Balb C/3T3 systems), and take appropriate action to facilitate the validation and acceptance of one or both of these systems.

3. A review of the need to conduct *in vivo* genotoxicity testing, especially for chemicals that are not intended for human consumption or direct exposure, and those for which indirect exposure is expected to be negligible.
4. The development of receptor-binding assays for non-genotoxic carcinogens.
5. The further development of QSAR and other computer-based approaches for mutagenicity and carcinogenicity.

9.7.2 Medium-term prospects

1. An appraisal of the usefulness of genomics and proteomics in the development of *in vitro* methods for genotoxicity testing.
2. The validation of receptor-binding assays for non-genotoxic carcinogens.
3. The validation of QSAR and other computer-based approaches for predicting genotoxicity and carcinogenicity.

9.7.3 Recommendations for research and development

1. Further research concerning: a) the known modes of action of non-genotoxic carcinogens; b) their relevance to human hazard; and c) the available and required non-animal testing strategies for detecting and characterising those carcinogens that are relevant.
2. The development of human cell-based cell transformation assays.

9.7.4 Other recommendations

1. Less emphasis should be placed on the development and validation of transgenic rodent models of carcinogenesis as assays for detecting carcinogens, since these are unlikely to be capable of identifying rodent and/or human carcinogens.
2. A way needs to be found, perhaps via the ICH process, and also via relevant regulatory agencies responsible for non-pharmaceutical products and for chemicals, to change the emphasis from predicting rodent carcinogenicity to predicting potential human carcinogenicity.
3. ECVAM should organise workshops on:
a) non-genotoxic carcinogens, in conjunction with the FRAME Toxicity Committee and the International Agency for Research on Cancer (IARC), to discuss and make recommendations for: i) the known modes of action of non-genotoxic carcinogens; ii) their relevance to human hazard; and iii) the available and required non-animal testing strategies for detecting and characterising those carcinogens that are relevant.
b) the use of genetically-engineered cell lines for predicting metabolism-mediated genotoxicity.

Figure 9.1.   A tiered testing strategy for genotoxicity.

1.   Use a validated (Q)SAR(s) for genotoxicity $\Rightarrow$ positive or negative result[1]
        $\Downarrow$

2.   Use validated *in vitro* tests for gene mutation
        and chromosome aberrations $\Rightarrow$ negative
                                                      $\Rightarrow$ assume no genotoxic potential[2]

                                                      $\Rightarrow$ positive
                                                      $\Rightarrow$ classify as genotoxic[2]

        $\Downarrow$
        ambiguous result[3]

        $\Downarrow$

3.   Conduct a short-term *in vivo* test[4] $\Rightarrow$ negative
                                                      $\Rightarrow$ assume no genotoxic potential.
                                                      Assess potential for non-genotoxic
                                                      carcinogenicity (Figure 9.2)

        $\Downarrow$
        positive result
        $\Downarrow$
        Classify as genotoxic
        (somatic cell mutagen) $\Rightarrow$ assume carcinogenic potential
        $\Downarrow$

4.   Use a combination of a validated
        germ cell mutagenicity test
        and biokinetic data[5] $\Rightarrow$ negative $\Rightarrow$ not a germ-cell mutagen
        $\Downarrow$
        positive or ambiguous result
        $\Downarrow$
        Classify as genotoxic (germ cell mutagen)

Footnotes

[1]The computer-based prediction should be used in conjunction with *in vitro* data, according to a weight-of-evidence approach.

[2]In most cases, the assessment of genotoxic potential would stop here. The need for further genotoxicity testing depends on usage and exposure levels.

[3]An ambiguous result could be followed up by *in vivo* testing, or by returning to step 2 and conducting different *in vitro* tests.

[4]Usually bone-marrow micronucleus or metaphase analysis, and/or unscheduled DNA synthesis in liver. The need for the short-term *in vivo* test should be assessed on a case-by-case basis.

[5]In general, somatic cell mutagens are assumed to be potential germ cell mutagens. This step is therefore optional, to provide a means of assessing whether a somatic cell mutagen is, or is not, a germ cell mutagen.

Figure 9.2.   A tiered testing strategy for carcinogenicity.

1.   Assess genotoxic potential
        (Figure 9.1) $\Rightarrow$ classify as genotoxic
                        $\Downarrow$
        non-genotoxic or ambiguous result
                        $\Downarrow$
2.   Use a validated (Q)SAR(s) for
        non-genotoxic carcinogencity $\Rightarrow$ positive $\Rightarrow$ classify as carcinogen
                        $\Downarrow$
        negative or ambiguous result, or no (Q)SAR available
                        $\Downarrow$
3.   Use validated *in vitro* models for
        non-genotoxic carcinogenicity[1] $\Rightarrow$ positive $\Rightarrow$ classify as carcinogen
                        $\Downarrow$
        negative or ambiguous result
                        $\Downarrow$
4.   Use an *in vivo* test[2] $\Rightarrow$ positive $\Rightarrow$ classify as carcinogen
                        $\Downarrow$
        negative or ambiguous result
                        $\Downarrow$
        unclassified

Footnotes

[1]Cell transformation assay, *in vitro* micronucleus assay, aneuploidy assay.
[2]Rodent bioassay or transgenic mouse model. The need for the *in vivo* test should be assessed on a case-by-case basis.

10. REPRODUCTIVE TOXICITY

10.1 Introduction

Reproductive toxicity refers to the adverse effects of a substance on any aspect of the reproductive cycle (Figure 10.1), including the impairment of reproductive function, the induction of adverse effects in the embryo, such as growth retardation, malformations, and death, and the induction of adverse post-natal effects. Traditional animal tests for reproductive toxicity include the pre-natal developmental toxicity study (OECD TG 414), the one-generation study (OECD TG 415), the two-generation study (OECD TG 416), the reproductive/developmental toxicity screening test (OECD TG 421), and the repeat-dose toxicity study combined with the reproductive/developmental toxicity study (OECD TG 422). In recent years, there has been a considerable increase in the number of short-term *in vivo* screening tests carried out in the context of reproductive toxicity testing, in order to detect so-called "endocrine disruptors" (Combes, 2000; Baker, 2001). This topic is discussed further in Chapter 11.

10.2 Scientific background

Since it is not possible to model the whole of the reproductive cycle *in vitro*, the components need to be studied individually. A number of useful and promising *in vitro* systems are currently available.

10.2.1 Gametogenesis

Gametes are derived from the primordial germ cells (PGCs), which enter the gonads during prenatal development. The PGCs arise at some distance from the presumptive gonads, to which they migrate and where they become established. The formation of the germ line is dependent upon the presence of the germ plasm, a cytoplasmic component that causes these cells to become distinct from the somatic cells. When the primordial germ cells are established in the gonad, they form stem cells that divide by mitosis to produce the supply of male and female gametes for reproduction. When they enter the gonads, the germ cells may associate with specific somatic cells that support, nurture and protect them. In the female, the somatic cells are the follicle cells. In the male, the comparable somatic cells are called Sertoli cells. During the proliferative phase, the germ cells are called oogonia in the ovary and spermatogonia in the testis. They act as a stem cell population that divides by mitosis to produce a lifetime supply of gametes for reproduction. The gonial cell divisions may be incomplete, with the result that the daughter cells maintain intercellular communication with one another via intercellular bridges. Successive incomplete divisions produce very large clones of interconnected cells. This intercellular communication may serve to synchronise the development of the conjoined cells. When the organism reaches maturity, germ cells acquire the ability to differentiate into functional gametes by meiosis, reducing the chromosome number from 2n to 1n.

10.2.2 Oogenesis

During the first meiotic division, the primary oocyte divides to produce one small polar body and one secondary oocyte. The latter will enter the second meiotic division

to produce the second polar body and the haploid ovum, which is the only functional sex cell to result. In most mammalian species, differentiation of the oocyte occurs during a protracted prophase of the first meiotic division. The resumption of meiosis, which occurs after the oocyte is fully grown, is called oocyte maturation.

10.2.3 Theca cells and granulosa cells

The ovaries provide three types of hormone: testosterone, oestrogen and progesterone. Testosterone is made by the theca cells that surround each follicle. Some of the testosterone is released into the bloodstream, but most of it is taken up by the granulosa cells that form the follicle and convert testosterone to oestrogen. As the follicles grow bigger, more and more oestrogen is produced and the level of oestrogen in the bloodstream rises. After ovulation, the empty follicle turns into the corpus luteum, which produces progesterone, which is important, since it prepares the uterus for implantation. Progesterone is only produced in the second half of the cycle, and its detection in the bloodstream provides proof that ovulation has occurred.

10.2.4 Spermatogenesis

In the male, meiosis precedes sex cell differentiation. A single spermatogonium enters the first meiotic division as a primary spermatocyte. This division produces two secondary spermatocytes, each of which divides to form two haploid spermatids. Each spermatid then differentiates (by a process called spermiogenesis) into a spermatozoon via specific structural and functional differentiation that enable the sperm to fertilise the egg. Consequently, four haploid sperm cells result from each diploid spermatogonium. The utilisation of all four haploid cells in the male is significant, since the testis must produce millions of sperm simultaneously.

10.2.5 Leydig cells and Sertoli cells

The testis can be distinguished into functional compartments, defined by their roles in the development of gametes or in steroid hormone production. The Leydig cells in the interstitial tissue of the testis are positioned so as to provide testosterone for the seminiferous tubules, which drive spermatogenesis, and to the peripheral circulation, which supplies testosterone to the rest of the body, to form and maintain male secondary sexual structures and characteristics.

Sertoli cells are non-proliferative, columnar epithelial cells extending from the basal lamina to the lumen of the seminiferous tubule. Adjacent Sertoli cells are united by band-like, occlusive (tight) junctions, which divide the seminiferous tubule into a basal compartment which extends from the basal lamina to the Sertoli cell tight junctions, and a luminal compartment which extends from the tight junctions to the lumen of the seminiferous tubule. Developing sperm cells occupy deep recesses in the lateral surfaces of Sertoli cells.

10.2.6 The blood-testis barrier

The tight junctional complexes between Sertoli cells create an impermeable wall that divides the seminiferous tubule into two compartments. The basal (outer) compartment is accessible to substances within the circulation (blood), whereas the adluminal (inner) compartment is not accessible to substances in the blood. As

gonocytes differentiate into specialised mature sperm cells, they move across the blood-testis barrier.

10.2.7 The placental barrier

The placenta is a materno-fetal vascular organ, which is formed during the implantation of the blastocyst, and is delivered with the fetus at birth. During pregnancy, it provides nutrition, gas exchange, waste removal, and endocrine and immune support for the developing fetus. The main function of the placenta is to permit the exchange of substances, rather than to act as a barrier. Most drugs and infectious agents readily pass across the barrier. The placenta produces several hormones, including human chorionic gonadotrophin, and the steroid hormones, progestin and oestrogen.

A general guideline for the administration of drugs during pregnancy is that whatever the mother receives, the fetus will also receive by diffusion across the placental barrier. However, the transport and metabolism of drugs in the placenta are poorly understood processes, because experimental studies with animals are difficult, and studies on humans are impossible. Furthermore, there are significant interspecies differences in placental function.

10.3 Current status of alternative methods for reproductive toxicity

An overview of alternative methods for reproductive toxicity testing is presented in Table 10.1. More-detailed information is available in ECVAM Workshop Report 2, and in a review article by Spielmann (1998). The following sections on the developing embryo and the placental barrier address *in vitro* methods for developmental toxicity, whereas the sections on primordial germ cells, germ cells, sperm motility and morphology, and Leydig and Sertoli cells cover other aspects of fertility.

10.3.1 The developing embryo

Birth defects and malformations in humans can be inherited or can result from exposure of the mother and embryo to drugs, occupational and environmental toxicants, or to other ingested substances, during critical and sensitive periods of development.

In the field of developmental toxicity, a variety of alternatives to animal testing are available. ECVAM has funded the prevalidation and validation of three different embryotoxicity tests: the whole embryo culture (WEC) test, the micromass (MM) test, and embryonic stem cell test (EST; Scholz *et al.*, 1999). For each assay, a prediction model has been developed, to classify the chemicals into three classes of embryotoxicity (non, weak/moderate, strong). The scientifically validity of the three methods was endorsed by the ESAC in October 2001: the EST and the WEC test were considered scientifically valid for distinguishing between non, weak/moderate and strong embryotoxins, whereas the MM test was considered scientifically valid for identifying strong embryotoxic chemicals.

In 1998, the US Environmental Protection Agency (EPA) asked ICCVAM to evaluate the Frog Embryo Teratogenesis Assay - *Xenopus* (FETAX), a 96-hour, whole-embryo

developmental toxicity test involving the anuran amphibian *Xenopus laevis*. In June 2001, an expert scientific committee concluded that FETAX is not sufficiently validated or optimised for regulatory use.

10.3.2 The placental barrier

Audus and his colleagues (University of Kansas, USA) are planning to investigate how drugs of abuse interact with the placenta *in vitro*, by using cell culture models of the placental barrier. The five *in vitro* models represent the early, middle, and terminal stages of pregnancy (Audus, 1999).

Another *in vitro* system, based on a fibre-based bioreactor design, uses a three-dimensional cell culture model which more closely mimics the trophoplast environment *in vivo*. An initial innoculum of trophoblast cells has been maintained in this system for over four weeks. Periodic removal and examination of cells demonstrated that the cells readily attached to the fibrous matrix and formed three-dimensional structures reminiscent of their *in vivo* architecture (e.g. villous-like structures). Electron microscopical examination of the fibrous matrix demonstrated that the trophoblast cells contained a complement of organelles similar to that found *in vivo*, indicating that the cells had achieved some degree of morphological differentiation. Metabolic studies of glucose consumption and lactate accumulation indicated that the culture reached a steady state within about 2.5 days, and the production of 17 β-oestradiol was maintained, indicating further metabolic activity.

10.3.3 Primordial germ cells

The *in vitro* maturation of germ cells in humans is a challenging process to model, because gametogenesis is a lengthy process, encompassing many complex cellular changes. PGCs migrate from the base of the allantois to the genital ridge. They proliferate both during migration and after their arrival, until the fetal gonads start to undergo sexual differentiation. PGCs then enter into the prophase of the first meiotic division in the ovary to become oocytes, while those in the testis become mitotically arrested to become prospermatogonia.

Regulation of the growth of mouse PGCs has been studied by culturing them on feeder cell layers. They show a limited period of proliferation *in vitro* before entering growth arrest, which corresponds with their developmental changes *in vivo*. However, in the presence of multiple growth signals, PGCs can restart rapid proliferation and transform into pluripotent embryonic germ cells (EGCs). Studies on ectopic germ cells and studies of reaggregate cultures have shown that both male and female PGCs undergo cell-autonomous entry into meiosis, and differentiate into oocytes if they are separated from the male gonadal environment (Nakatsuji & Chuma, 2001). Further research is necessary to determine whether some of the existing PGC lines (Klemm *et al.*, 2001) still retain the characteristics of PGCs (Shamblott *et al.*, 2001).

In addition, two-dimensional dispersed culture systems have been developed, in which the transition from the mitotic PGCs into the leptotene stage of the first meiotic division can be analysed. Entry into meiosis seems to be programmed in PGCs before they reach the genital ridges, and unless it is inhibited by putative signals from the testicular somatic cells.

10.3.4 Spermatogonia and oocytes

An ambitious project was funded by the EC in the Fourth Framework Programme. The aim of the study was to establish germ cell culture systems and relevant markers for toxicological endpoints. Primary oocytes and spermatogonial cells were cultured *in vitro* and exposed to germ cell toxicants with cytotoxicity, apoptosis and selected gene expression as toxicological endpoints. It was demonstrated that the *in vitro* model is more sensitive than the *in vivo* system. Unfortunately, no standard operating procedure or INVITTOX protocol was developed, so the system cannot be considered to be ready for prevalidation (Cuzin & Rassaoulzadegan, 1998, De Felici, 1998, De Felici *et al.*, 1998).

Another project on the production of oocytes (entitled *In vitro production of high quality mammalian oocytes for biotechnology, assisted reproductions, breeding and toxicology-teratology purposes*) is currently under way. A positive outcome for this study could lead to a source of oocytes that could be used for toxicological studies.

10.3.5 Sperm motility and morphology

Gametogenesis is a complex biological process that is sensitive to environmental insult, e.g. from chemicals. Chemical effects on germ cells and their maturation can inhibit fertility and may have negative effects on the development of the offspring. Mutagens, for example, produce heritable gene mutations, and heritable structural and numerical chromosome aberrations in germ cells. The consequences of germ cell mutation for subsequent generations include genetically-determined phenotypic alterations without signs of illness, reduction in fertility, and embryonic or perinatal death, congenital malformations with varying degrees of severity, and genetic diseases with varying degrees of health impairment.

Semen analysis permits the detection of the effects of chemicals on post-testicular stages. This kind of analysis is important in two ways: a) changes in sperm or seminal content may be indicative of adverse effects on the male reproductive system, with possible implications for fertility potential; and b) defects in sperm DNA or chromosomes may be associated with subsequent changes in viability during embryonic development, with health risks to the offspring.

Several techniques are available for monitoring sperm motility, motion, morphology, head morphology, chromatin structure and various other aspects of semen composition. Some of these analyses can be automated (for example, by videography). A practical system has been well developed and an INVITTOX protocol is available (Slott *et al.*, 1993, Hinsch *et al.*, 1997, Vetter *et al.*, 1998).

10.3.6 Leydig cells and Sertoli cells

During the last few years, there have been some claims that there is evidence of a deterioration in mammalian (including human) male reproductive health, leading to suggestions that a number of chemicals found in the environment act as endocrine disrupters. In addition, a wide range of compounds have been reported to be toxic to one of the pivotal cell types involved in spermatogenesis, namely, the Leydig cell.

A test system for detecting the adverse effects of chemicals and environmental pollutants on male fertility was developed in the EC's Fourth Framework Programme. Intercellular interactions in the testis have been modelled by using co-cultures of Leydig and Sertoli cells (Fielder *et al.*, 1997). Leydig cell lines showing constitutive activity in steroidogenesis have been established, and several toxicants have been tested with regard to their effects on steroidogenesis, cell proliferation and cell viability (Cooke, 1992, 1998). ECVAM is planning to follow up this work by organising a prevalidation study on the use of Leydig cell lines.

Additional research on the blood-testis barrier is required, to establish and characterise the chemical classes that are able to cross the barrier and affect germ cell maturation. An understanding of the mechanisms of this barrier could help to decrease the number of possible germ cell mutagens that have to be tested (Janecke, 1992; Kumi-Diaku *et al.*, 1999; Chen *et al.*, 2001; Harrison *et al.*, 2001).

10.4 Projects funded by the European Commission

The projects that have been funded by DG Research are summarised in Table 10.2. ECVAM plans to contact the coordinators of some of these projects, to discuss the possibility of following them up in terms of ECVAM prevalidation studies.

ECVAM has funded, or is funding, the following studies:

1. The ECVAM prevalidation and validation studies on three alternative tests for embryotoxicity.
2. An investigation on reporter gene expression in embryonic stem cell-derived embryoid bodies as an endpoint for identifying potential embryotoxic/teratogenic substances.
3. *In vitro* methods for metal-induced infertility and spermiotoxicity

10.5 Reproductive toxicity: summary, conclusions and recommendations

Since it is not possible to model the whole of the reproductive cycle *in vitro* with one approach, the components need to be studied individually and then integrated into testing strategy.

For the assessment of embryotoxic potential, the EST, the WEC and the MM tests can be used immediately. However, the EST is the only one of the three which could have a relatively high throughput, and which would not involve the killing of large numbers of pregnant animals.

It will take time to develop and validate a battery of alternative tests that can cover the various aspects of the reproductive cycle, so animal tests will continue to be required for the foreseeable future, at least for certain aspects of reproductive toxicity testing. Therefore, a short-term priority should be to refine existing animal tests for reproductive toxicity.

10.5.1 Short--term prospects (developmental toxicity)

1. The transfer of the EST to interested laboratories.

2. The development of test guidelines based on the EST, WEC and MM tests, to be submitted for regulatory acceptance.
3. The refinement of the EST with target cell-specific endpoints and human embryonic stem cells.

10.5.2 Medium-term prospects (developmental toxicity)

A catch-up validation study, to assess the validity of the refined EST.

10.5.3 Long-term prospects (developmental toxicity)

The development and evaluation of a testing strategy which covers not only malformations, but also other manifestations of developmental toxicity, such as growth retardation and embryolethality.

10.5.4 Short-term prospects (fertility)

1. The identification of the most predictive toxicological endpoints for use with semen analysis.
2. The prevalidation of a Leydig cell line test.

10.5.5 Medium-term prospects (fertility)

The prevalidation of methods for metal-induced infertility and spermiotoxicity.

10.5.6 Long-term prospects (fertility)

1. The development and evaluation of a testing strategy covering all essential aspects of the male and female reproductive cycles.
2. The validation of a Leydig cell line test.
3. The validation of methods for metal-induced infertility and spermiotoxicity.

10.5.7 Recommendations for research and development

1. An evaluation of the use of primary cultures of spermatogonia and oocytes.
2. Research on the use of PGCs and/or PGC lines for identifying germ cell mutagens.
3. The further development of granulosa cell and theca cell assays as toxicological screening systems.
4. An evaluation of the use of Sertoli cell lines (including Sertoli cell co-cultures) for toxicological purposes.
5. Further research on the blood-testis barrier, to define its permeability characteristics and the chemical classes that can cross it.

10.5.8 Other recommendations

ECVAM should organise workshops on:

a) embryotoxicity, to define the areas of application of the scientifically-validated EST; and
b) the use of hormone-producing cells for predicting the adverse effects of chemicals on fertility.

Figure 10.1.   The mammalian reproductive cycle.
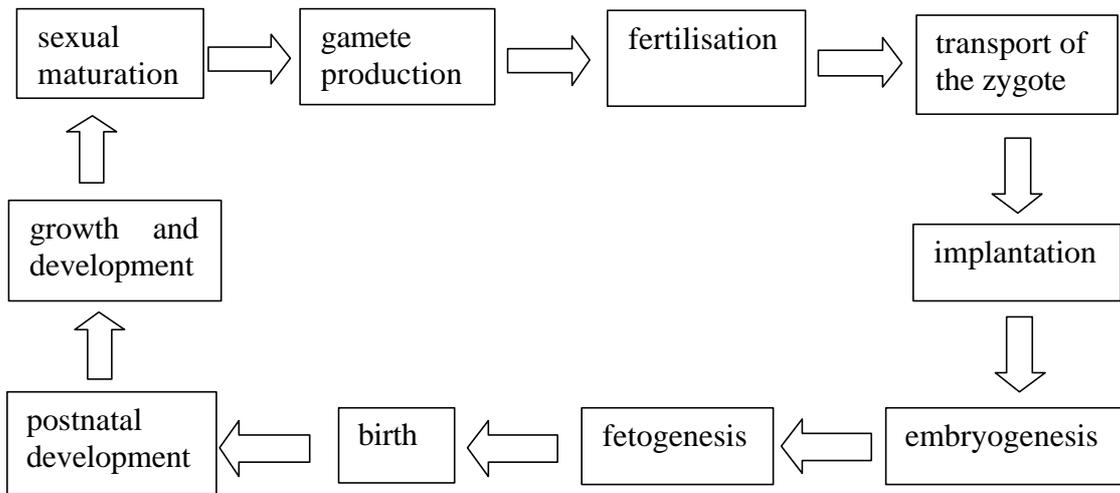
Footnote to Figure 10.1

The figure excludes behavioural aspects and the fetal/maternal interactions.

Table 10.1. An overview of *in vitro* methods for reproductive toxicity.

Target cells/organs	Test System	Endpoint	Applicability	Formal status
PGC	Primary cells PCG/EGC cell lines	Viability, sister chromatid exchanges, mutagenicity	Germ cell mutagens	Further development required
Spermatogonia	Primary cells Cell lines (GC-1 spg)	Specific gene expression	Chemicals affecting fertility	Further development required
Spermatocytes	Primary cells	Differential display	Chemicals affecting fertility	Further development required
Semen	Primary cells	Sperm motility, motion, morphology	Chemicals acting postmeiotically	Ready for prevalidation
Sertoli cells	Cell lines (15P-1, TM 4)	Cell survival/growth	Chemicals destroying the blood-testis barrier and the support of germ cell maturation	Further development required
Leydig cells	Leydig tumor cell lines (MA10; RC2, TM 3, MLTC) and genetically engineered HEK 293 cells with LH receptors	Cytotoxicity, proliferation, apoptosis, steroidogenesis, western / northern blotting, steroid dehydrogenesis	Chemicals interfering with endocrine function (endocrine disrupters)	Readiness for prevalidation needs to be assessed
Fetal oocytes	Primary cells	Viability, specific gene expression	Chemicals affecting fertility	Further development required

Theca cells	Primary cells	Not developed	Chemicals interfering with endocrine function (endocrine disrupters)	Further development required
Granulosa cells	Primary cells	Not developed	Chemicals destroying the support of germ cell maturation	Further development required
Blood-testis barrier	Sertoli cell lines (TM 4)	Not developed	Chemicals crossing the barrier	To be evaluated by ECVAM
Placenta and placenta barrier	Trophoblast cell lines (BT1)	Not developed	Chemicals inducing placental toxicity	To be evaluated by ECVAM
Developing embryo	WEC, MM, EST	Morphological endpoints and cytotoxicity	Direct acting embryotoxic/teratogenic compounds	Scientifically validated and endorsed by the ESAC

Table 10.2.   Projects on reproductive toxicity funded by the European Commission (DG Research).

1	"Development and evaluation of Leydig cell lines as *in vitro* models for toxicological testing" Project Reference: BIO 4 CT 972022 Start date: 01-10-1997 End date: 30-09-2000 Duration: 36 months
2	"Development of 3D *in vitro* models of human tissues for pharmaco-toxicological applications" Project Reference: BIO4972148 Start date: 01-09-1997 End date: 31-08-2000 Duration: 36 months
3	"Identification of critical rat testicular genes altered after fetal androgenic disruption by flutamide: use of DNA microarray approach" Project Reference: QLK4-2000-00684 Start date: 01-01-2001 End date: 31-13-2003 Duration: 36 months
4	"*In vitro* production of high quality mammalian oocytes for biotechnology, assisted reproduction, breeding and toxicology-teratology purposes" Project Reference: QLK3-1999-00104 Start date: End date: Duration: 36 months
5	"Development of *in vitro* mammalian germ cell culture systems and genetic markers for reproductive pharmaco-toxicology" Project Reference: BIO4-CT96-0183 Start date: 1-10-1996 End date: 30-09-1999 Duration: 36 months

6	"Development of a new test of developmental toxicology" Project Reference: Start date: 01-11-1993 End date: 31-10-1996 Duration: 36 months
7	"The use of gene transfer technology in conjunction with primary and clonal culture for the establishment of predictive *in vitro* screening assays for teratogenic potential : development, validation and transfer to industry" Project Reference: BIO4-CT96-0183 Start date: 01-09-1993 End date: 31-08-1996 Duration: 36 months

11. ENDOCRINE DISRUPTION IN HUMANS

11.1 Introduction

11.1.1 Concerns about possible wildlife and human health effects

Many *in vivo* and *in vitro* tests have been proposed for screening chemicals for endocrine disrupting activity, and several regulatory agencies, including the US Environmental Protection Agency (EPA), have recommended strategies involving tiered testing schemes, in response to concerns about the possible adverse effects of such chemicals on the hormonal systems of wild-life and humans. Such chemicals have been called endocrine disruptors (EDs) and are alleged to mimic (agonise) or block (antagonise) the effects of endogenous sex steroid hormones. Reported adverse human health effects associated with exposure to them include deterioration in semen quality, various defects on gonadal development, and testicular, prostate, uterine and breast cancer, as well as hypospadias, endometriosis and effects on the thyroid gland (see Zacharewski, 1998; Combes, 2000a).

11.1.2 US legislation

Legislation in the USA requires the testing of chemicals for endocrine activity (the *Food Quality Protection Act* of 1996, Public Law 104-170, and the *Safe Drinking Water Act Amendments* of 1996, Public Law 104-182; Anon, 2001a). Other regions of the world, and especially Japan and Europe, are also introducing legislation for ED testing (for example, Anon, 2001b).

A major problem with investigating the possible modulation of the hormonal system by chemicals is that endogenous sex steroid hormones (such as oestrogen, testosterone and progesterone, characterised by the presence of a 4-ringed fused cyclopenteno-phenanthrene nucleus possessing high lipid solubility) control normal development.

11.1.3 Definition of endocrine disruptor

Various definitions of EDs have been proposed. For example, the International Programme on Chemical Safety (IPCS) and the World Health Organisation (WHO) define an endocrine disrupting chemical as *"an exogenous substance or mixture that alters function(s) of the endocrine system and causes adverse health effects in an intact organism, or its progeny, or (sub)populations."* This is the definition that is used in this chapter, although the reader is referred to Harvey *et al.* (1999) for further discussion.

11.1.4 Types of endocrine disruptors

A wide diversity of chemicals can exert deleterious effects on the endocrine system, including synthetic chemicals (e.g. polychlorinated biphenyls, phthalates, organochlorines, polychlorinated dibenzo-*p*-dioxins, dibenzofuran, and a range of synthetic steroids) and naturally-occurring chemicals (e.g. flavonoids, lignans, sterols, aromatic amines and indole-3-carbinol).

11.2 Scientific background

11.2.1 The endocrine system

The endocrine system comprises many organs producing a number of different hormones, that are part of a complex feed-back regulatory system (see Crain *et al.,* 2000, for a review).

Endogenous hormones act primarily by binding specifically to an extensive family of nuclear receptors (steroid hormone receptors or SHRs; Warner *et al*., 1999). SHRs have structural features in common, which include a central highly-conserved DNA-binding domain that targets the receptor to specific DNA sequences (hormone response elements, HREs). The terminal portion of the receptor includes a ligand-binding domain that interacts directly with the hormone. This domain contains a hormone-dependent transcriptional activation region. Binding of the steroid hormone to the receptor protein causes it to undergo a conformational change, resulting in binding of the DNA to the HRE, to form a complex that triggers or suppresses the transcription of a specific set of genes (Wiegel, 1996; Combes, 2000a).

11.2.2 The activity of endocrine disruptors

EDs act by a diversity of mechanisms, including receptor binding, altered post-receptor activation, steroidogenesis (modulation of hormone synthesis), hormone storage and clearance, and perturbation of homeostasis. The immediate effects of this activity, depending on the nature of the ED, are: a) oestrogen agonism or antagonism; b) androgen agonism or antagonism; c) progesterone agonism or antagonism; d) suppression or induction of levels of endogenous hormones; e) and other effects, such as the release of follicle stimulating hormone (FSH), luteinising hormone (LH) or prolactin.

The oestrogen receptor (ER) exists in at least 2 subtypes, α and β, each encoded by a separate gene. The ER subtypes differ in their tissue distribution and relative ligand-binding affinities for the same hormones (a fact that could account for differential hormonal effects according to tissue).

11.3 Tests for endocrine disruptors

Testing for EDs is a new area of toxicology, and testing strategies are currently being developed. Some 56 different tests have been identified for detecting EDs (Holmes *et al*., 1998; Anon, 2000; summarised in Tables 11.1 and 11.2). These methods have been combined into tiered testing strategies comprising *in vitro* and *in vivo* assays.

11.3.1 In vivo assays

The rodent uterotrophic assay, one of the original methods developed for studying oestrogenicity, is still regarded as one of the most useful indicators of endocrine disruption. The test directly measures stimulation of mitotic activity caused by binding to the oestrogen receptor, in tissues of the female reproductive tract. Another animal test was devised by Hershberger and colleagues, to detect the ability of a chemical to elicit agonistic or antagonistic effects by binding to the androgen receptor in male rodents. It is necessary to eliminate any interfering effects from endogenous

sources of oestrogen and androgen, respectively, in these tests, either by using ovariectomised and castrated animals, or by using immature ones.

Many other *in vivo* tests have been, or are being, developed and some of these are summarised in Table 11.1.

11.3.2 Drawbacks of in vivo tests

There are four major problems with the available animal tests for EDs: a) lack of reproducibility; b) insufficient validation; c) difficulties with applying standard validation criteria; and d) concerns about interpreting the resulting hazard information for risk assessment. EDs can exert contrasting activities according to the dose administered and the nature of the target cells. Many EDs exert very weak effects, and there are problems with intra-laboratory and inter-laboratory reproducibility. Also, the interpretation of data can be confounded by a number of factors, including: a) the oestrogenic activity of components of the normal animal diet; b) the use of different species and strains; and c) differences in animal housing and husbandry.

The above problems, together with the experimental procedures required, contribute to severe animal welfare concerns arising from the use of animal tests. Some of the assays require the use of considerable numbers of animals in time-consuming experiments, which can involve complicated procedures, possibly including surgery, implantation and repeated injections. Also, considerable animal wastage could occur, when conducting assays that involve animals of only one sex.

11.3.3 In vitro assays

Several *in vitro* methods are available as screens for detecting EDs, which are designed to be sensitive and rapid, facilitating the testing of large numbers of chemicals (see Combes, 2000a, and Baker, 2001, for reviews). The main tests involve subcellular hormone receptor ligand-binding and the induction of proliferation (mitogenesis) in hormone-responsive mammalian cell lines, as well as transactivation systems in yeast and mammalian cell lines (Table 11.2).

(Subcellular) hormone receptor ligand-binding assays

These assays involve assessing the molecular binding of the hormone to isolated and purified receptor protein. Competitive ligand-binding assays, whether alone or combined with reporter gene expression, are designed to detect chemicals which directly interact with the endocrine receptor. Similar cell-free binding assays involving the androgen receptor (AR) have been developed (Kelce & Gray, 1999).

153

Hormone-responsive mammalian cell proliferation

The most widely used test is based on the induction of mitogenesis in oestrogen-responding cells, particularly in the human breast cancer cell line, MCF-7. MCF-7 cells express ER, AR, as well as receptors for progesterone, glucocorticoid, vitamin D and retinoic acid. The induction of proliferation in these cells is believed to result from the binding of oestrogenic substances to the ER within the cells. Several other *in vitro* assays have also been developed which involve the use of diverse human and animal tissues and primary cell cultures.

Hormone-responsive mammalian cell proliferation

In vitro assays for measuring steroidogenesis include the use of primary cultures of Leydig cells of the testis (see also Chapter 10). These cells are primarily responsible for steroidogenesis and are susceptible to various androgen-binding chemicals, such as flutamide. Ovarian steroidogenesis can be measured *in vitro*; for example, Johannsen *et al.* (2002) have described a method based on measuring glucocorticoid steroidogenesis in a human adrenocortical carcinoma cell line (H295R).

Hormone-sensitive transcription of reporter genes

These tests use genetically-engineered mammalian cells or strains of yeast (*Saccharomyces cerevisiae*) expressing human or mammalian ER nuclear receptors, in conjunction with the respective response elements linked to promoter regions for a reporter gene, together with the reporter gene itself. Receptor activation is detected as a stimulation of reporter gene expression. Several different receptor-response element-reporter gene combinations have been utilised in yeast (e.g. ERE-gal1-lacZ for detection of oestrogenicity via β-galactosidase production, and in mammalian cell transactivation systems (e.g. ERE-luciferase or ERE-CAT (chloramphenicol transferase). Strains of mammalian cells have also been constructed for the detection of chemicals with prostagenic and androgenic activities via interaction with the appropriate receptors.

Genomic analysis is also beginning to be applied to the screening of chemicals for ED activity, by developing microarrays containing DNA sequences for ER and HREs (J. Kanno, personal communication).

11.3.4 Drawbacks of in vitro assays

Apart from the limitations inherent in using all *in vitro* systems, several EDs are known to bind strongly to serum proteins. Such binding can occur with components of culture media, which results in variable *in vitro* data, as well as in differences in responses obtained in animal and non-animal assays. This phenomenon has been alleviated by removing serum from the medium, but this raises the problem that the *in vitro* model does not then account for serum binding *in vivo*. *In vitro* assays require the addition of an exogenous metabolising system. However, for ED testing, this has been considered to be confounded by the binding of the test chemical and its metabolites to components of the enzyme fractions used. However, Hashimoto *et al.* (2001) have shown that it is feasible to detect the effects of exogenous metabolism on ED activity *in vitro*.

11.3.5 (Q)SAR studies

Several studies have been conducted to identify the structural features of steroid molecules necessary for efficient binding to nuclear hormone receptors, such as the ER (see Blair *et al.*, 2000). Many oestrogens contain one or more phenolic OH groups on a small lipophilic molecule of about 200-300 Daltons. Removal of these moeities normally substantially reduces binding efficiency. One obvious reason why many chemicals seem to possess oestrogenicity is that the phenolic OH group is a commonly-used substitutent in synthetic chemistry. It should also be remembered that this chemical moeity is also found widely in the body, arising from the first-pass metabolism of endogenous and xenobiotic chemicals due to mixed-function oxidase activity.

Several computer-generated models of ligand-ER interactions have been produced, for investigating the mechanisms of binding and for predicting binding to the ER. This work is limited by the availability of crystal structures of the various SHRs. The reader is referred to Jacobs & Lewis (2002) for more information.

Several pharmacophore and QSAR models have been generated for predicting the binding affinity of chemicals to the ER (see Combes 2000a,b). However, most commercially-available expert systems lack extensive rulebases for detecting EDs. The MULTICASE system identifies a six-Angstrom unit spacer biophore as being associated with ED activity (see Combes, 2000b).

Hong *et al.* (2002) developed a tree-based model to evaluate 58,000 chemicals for binding to the ER, by using physicochemical information and SARs derived from a training set of chemicals with known binding affinities for the rat ER. The model also uses rejection (exclusion) filters, the first of which is a molecular weight range of <94 or >1,000. The second (inclusion) filter was the need for a ring structure and three structural alerts: a) the steroid ring; b) the phenolic ring; and c) the double ring structure of the diethylstilboestrol skeleton. Filtering was also based on log P, positive surface charge area (incorporating molecular shape and electronic distribution) and the breadth of the molecule. This computational approach resulted in a low false negative rate and predicted that 80% of the 58,000 chemicals (i.e. some 46,000) would show negligible binding to the ER.

11.4 Existing testing strategies for endocrine disruptors

11.4.1 Activities of the US EPA

The US EPA has been charged by the US Congress with the task of developing a screening strategy involving the use of validated tests, or other scientifically valid information. The EPA therefore established an Endocrine Disruptors Screening and Testing Advisory Committee (EDSTAC) to consider possible testing strategies for EDs, and to propose tiered testing schemes. EDSTAC proposed a testing strategy based on the combined use of different tests (Kavlock *et al.*, 1996; Shelby *et al.*, 1996). The committee suggested a strategy comprising an initial priority setting stage, followed by a Tier I screening phase, to be succeeded by a Tier II testing battery for detecting and characterising EDs (Figure 11.1). Tier I tests are intended to identify substances with endocrine-disrupting potential for further testing, and Tier II tests (all

in vivo tests) are designed to identify adverse effects and establish dose-response relationships for hazard and risk assessment.

Priority setting rank-orders the most important chemicals for more resource-intensive and more-costly Tier I evaluations. This entails the organisation of data on chemicals according to exposure, biological activity and statutory criteria. All of the information obtained is being entered into a database (the Endocrine Disruptor Priority Setting Database [EDPSD]). After prioritisation, the highest priority chemicals will be tested first. However, developing the EDPSD has been exacerbated by a general lack of information and delays in developing and using receptor-binding tests to provide some of this information. More than 87,000 chemicals were initially selected for evaluation, many of which were polymers or were otherwise considered unlikely to be active, leaving about 58,000 chemicals to be tested in Tier I.

11.4.2 Validation of test methods

11.4.2.1 The EDMVS

A Federal Advisory Committee of the EPA, the Endocrine Disruptor Methods Validation Subcommittee (EDMVS), has been established to advise the EPA on the development and validation of assays for ED testing.

11.4.2.2 Time-table for implementing the test strategy

The original EPA time-table for implementing a testing strategy for testing for Eds was: a) priority setting and validation of Tier I tests (2000-2003); b) validation of Tier II tests (2000-2005); and c) commencement of chemical screening (2003-2005).

11.4.2.3 Involvement of the OECD

An OECD Working Group on Endocrine Disruptor Testing and Assessment (EDTA) has been established to identify tests for validation, and a separate validation management committee is overseeing the validation of some of the *in vivo* tests identified by EDSTAC, namely the uterotrophic and Hershberger assays, and an enhanced repeat-dose toxicity protocol (OECD TG 407).

11.4.2.4 Progress with validation

The validation of tests for EDs poses a number of problems (see Ashby, 2000; Combes, 2000a). A series of validation studies has been conducted on the uterotrophic assay, and these have established that it can reliably detect both strong and weak oestrogenic substances (Kanno *et al.*, 2001; Combes 2002). However, the predictivity cannot be fully assessed due to a shortage of chemicals lacking activity. As yet, the uterotrophic assay cannot be recommended for screening anti-oestrogenic substances for regulatory purposes, because insufficient numbers of anti-oestrogens and laboratories have been involved in the validation studies conducted so far.

The validation studies for the other two assays are at various stages of progress. Phase 1 of the validation of the Hershberger assay was delayed due to a need to resolve several fundamental protocol issues (e.g. how to weigh testicular material).

However, a protocol has now been agreed, and phase 2 of the study is underway. With regard to the Enhanced 407 test, a prevalidation study has been conducted to investigate inter-laboratory transferability of the modified protocol, and to ascertain the usefulness of the additional endpoints in the test. It is intended that the results for the same chemicals tested in the Uterotrophic, Hershberger and Enhanced 407 assays will be compared.

Studies with the uterotrophic and Hershberger assays have shown that there is essentially no need to use ovariectomised or castrated animals (Ashby & Lefevre, 2001; Ashby *et al*., 2001), and this should be taken into account when OECD test guidelines are being developed for the tests. Also, some investigators have reported low-dose (<5mg/kg bwt) effects. This is a complex issue, although the current view of the NIEHS Low-Dose Peer Review Group is that the existence of low dose effects requires confirmation, since there have been many problems in repeating such results. Resolving the issue has been exacerbated by the technical difficulty of detecting very subtle changes, when these can be affected by animal strains, housing, handling, husbandry and feed. Thus, at the present time, the need for the inclusion of more dose groups in animal tests for EDs has not been scientifically established.

11.5 Discussion

There is considerable controversy concerning the potential adverse effects of EDs on human health. There is a very urgent need for careful and realistic assessments of likely exposure and for appropriate epidemiological analyses of susceptible groups. Such groups include those manufacturing and taking the contraceptive pill before, and unknowingly during, pregnancy (the latter for *in utero* exposure), and those involved in the manufacture and/or use of substances with well-known ED activity.

There has also been debate about the relative uses of non-animal and animal approaches for the detection of EDs. It is widely felt that a fully functional hormonal system is required to fully investigate whether a substance acts as an endocrine disruptor. Thus, until it is possible for the hormone system to be modelled outside the body, *in vivo* tests will be required to confirm predictions of ED activity. Thus, a positive result in a non-animal system based on mechanisms of action of EDs can indicate endocrine-disrupting potential. However, such activity detected in rodents can only indicate a potential for the substance to act as an endocrine disruptor in humans, because of considerable inter-species differences.

A major problem with developing and validating tests for EDs relates to the absence of general agreement on any definitive test data that can be used to provide a high level of certainty that a chemical is not an ED. In view of the diverse and complex effects of ED activity, the definitive test for ED activity is considered by some to be a multi-generation reproductive study. However, such tests are costly and time-consuming, and there is a paucity of data from them. The relevance of other tests, such as the uterotrophic assay, is limited by the difficulty of defining a PM for the assay. In other words, it is not clear how evidence that a chemical binds to the ER in a rodent should be interpreted, when it is known that endogenous hormones bind to

their respective receptors as part of developmental processes, including uterine development.

It is becoming clear that the uterotrophic and Hershberger assays can only be used to indicate that a substance is a *potential* ED in humans, and that they should be used in ways analogous to the use of the bone-marrow micronucleus assay in genotoxicity testing (Chapter 9), albeit with respect to a different target organ. The relevance to human health of a positive result in either of these *in vivo* assays is therefore not necessarily any more significant than a demonstration of receptor binding *in vitro*, coupled with evidence of associated transcriptional activation, especially if these latter endpoints were applied in systems incorporating both phase I and phase II metabolism, and were interpreted in conjunction with biokinetic and ADME information. It is therefore urgent that such *in vitro* assays are validated for use as the first part of a sequential testing strategy, i.e. to indicate potential for ED activity.

11.6 Endocrine disruptors: summary, conclusions and recommendations

The mechanisms whereby EDs act are complex and diverse, and this has hindered efforts to develop reliable and relevant animal and non-animal tests for their detection. The emergence of relevant and reliable animal tests has also been hindered by the fact that few, if any, of the available animal test methods have been designed with the specific purpose of detecting EDs. They are, in fact, modifications of existing animal tests. The use of animal tests for EDs poses several important welfare problems, and there are complications due to the possible existence of a low-dose effect. The validation of currently-available non-animal approaches should be made a priority, so that they can be used for compound prioritisation, and incorporated into tiered testing strategies in such a way that they reduce or replace further testing in animals. In addition, there is much research to be undertaken to assess the need for non-animal assays to cover the full spectrum of possible modes of action of EDs, and to provide definitive evidence of whether a substance does or does not interact with the hormonal systems of humans and animals.

11.6.1 Short-term prospects

The validation of (Q)SAR models and cell culture systems for predicting receptor binding.

11.6.2 Recommendations for research and development

1. Investigations on the use of microarrays for screening for receptor binding and gene induction, coupled with the search for further relevant receptors.
2. Research into the use of *in vitro* receptor binding assays in conjunction with biokinetic modelling.
3. The assessment of the potential use of biomarkers of exposure and effect for EDs.
4. Investigations on the suitability of *in vitro* systems for yielding consistent data regarding the potential interactive effects of chemical mixtures of EDs.
5. Further research on the basic mechanisms of ED action, especially via non-receptor pathways, and the development of *in vitro* models for such processes.
6. The development of appropriate *in vitro* methods for measuring steroidogenesis in males and females and for investigating endocrine function.

7. The development of approaches for detecting ED activity by using exogenous metabolising systems and cell culture systems, comprising transgenic cell lines with relevant hormone-receptor response elements, reporter gene sequences, and genes for phase I and phase II metabolism.
8. The development and application of immortalised fetal and pre-pubertal cell lines.
9. The development of an integrated testing strategy for EDs, based on the maximum use of non-animal approaches (Figure 11.2).

11.6.3 Recommendations to specific organisations

1. The organisation of an ECVAM workshop, to discuss the possibilities of using *in vitro* systems for screening for EDs, and to define the roles of such tests for this purpose. The workshop should include experts on endocrine disruption, *in vitro* toxicologists, clinical endocrinologists, ecologists, chemists, experts in (Q)SAR, and regulators. The workshop should focus on the limitations of *in vitro* approaches, and the need for them in relation to *in vivo* tests, bearing in mind that so far, no uniquely *in vivo* positive chemicals have been identified.
2. The appropriate regulatory authorities should consider the scientific merits of a test battery for EDs, either to supplement the existing reproductive toxicity test package, or to be applied as a separate set of tests specifically for ED activity.

Figure 11.1 A testing strategy for EDs, as recommended by the EDSTAC of the EPA.

```
                    ┌─────────────────────────┐
                    │       COMPOUND          │
                    │    PRIORITISATION        │
                    │    (use of EDPSD)*       │
                    └─────────────────────────┘
```

TIER I SCREENING ASSAYS	*TIER I ALTERNATIVE/EXTRA ASSAYS*
for preliminary hazard identification	AROMATASE (EPA)
RECEPTOR BINDING (ICCVAM for EPA)+	PUBERTAL MALE/FEMALE (EPA)
REPORTER GENE (ICCVAM for EPA)	ADULT 14 day INTACT MALE (INDUSTRY)
UTEROTROPHIC (OECD)	*IN UTERO* LACTATION (EPA)
HERSHBERGER (OECD)	
PUBERTAL FEMALE (EPA for OECD)	
FROG THYROID (EPA for OECD)	
FISH REPRODUCTION (EPA for OECD)	
STEROIDOGENESIS (EPA)	

```
                    ┌─────────────────────────┐
                    │      TIER II ASSAYS      │
```

TIER II ASSAYS

For further hazard identification

MAMMALIAN 2 GENERATION (EPA/OECD)
AVIAN 2 GENERATION (EPA for OECD)
CHRONIC AMPHIBIAN (EPA for OECD)
CHRONIC FISH (EPA for OECD)
CHRONIC INVERTEBRATE (EPA for OECD)

*	Endocrine Disruptor Priority Setting Database
+	Indicates organisation leading the development, validation and assessment of the assay

Figure 11.2. Elements of an integrated testing scheme for endocrine disruptors.

IN SILICO
(Q)SAR

↓

SUBCELLULAR
RECEPTOR BINDING
(to: oestrogen, androgen, thyroid and glucocorticoid receptors)
genomic analysis

↓

IN VITRO
CELL PROLIFERATION
GENE EXPRESSION/TRANSACTIVATION ASSAYS
(+/- exogenous metabolism & development of cell lines expressing
phase I and II metabolism and hormone receptors)

↓

STEROIDOGENESIS
*(aromatase & 5a-reductase inhibition; non-specific cell-
based tests)*
LIVER ENZYME INDUCTION

↓

IN VIVO
RECEPTOR BINDING
(uterotrophic & Hershberger tests)

↓

OTHER ENDPOINTS
*(enhanced 407; in utero lactation,
pubertal studies; multigeneration study)*
steroidogenesis
endocrine function
CNS gonadal axis effects
reduced fertility
fecundity
litter size
growth rates
mating indices
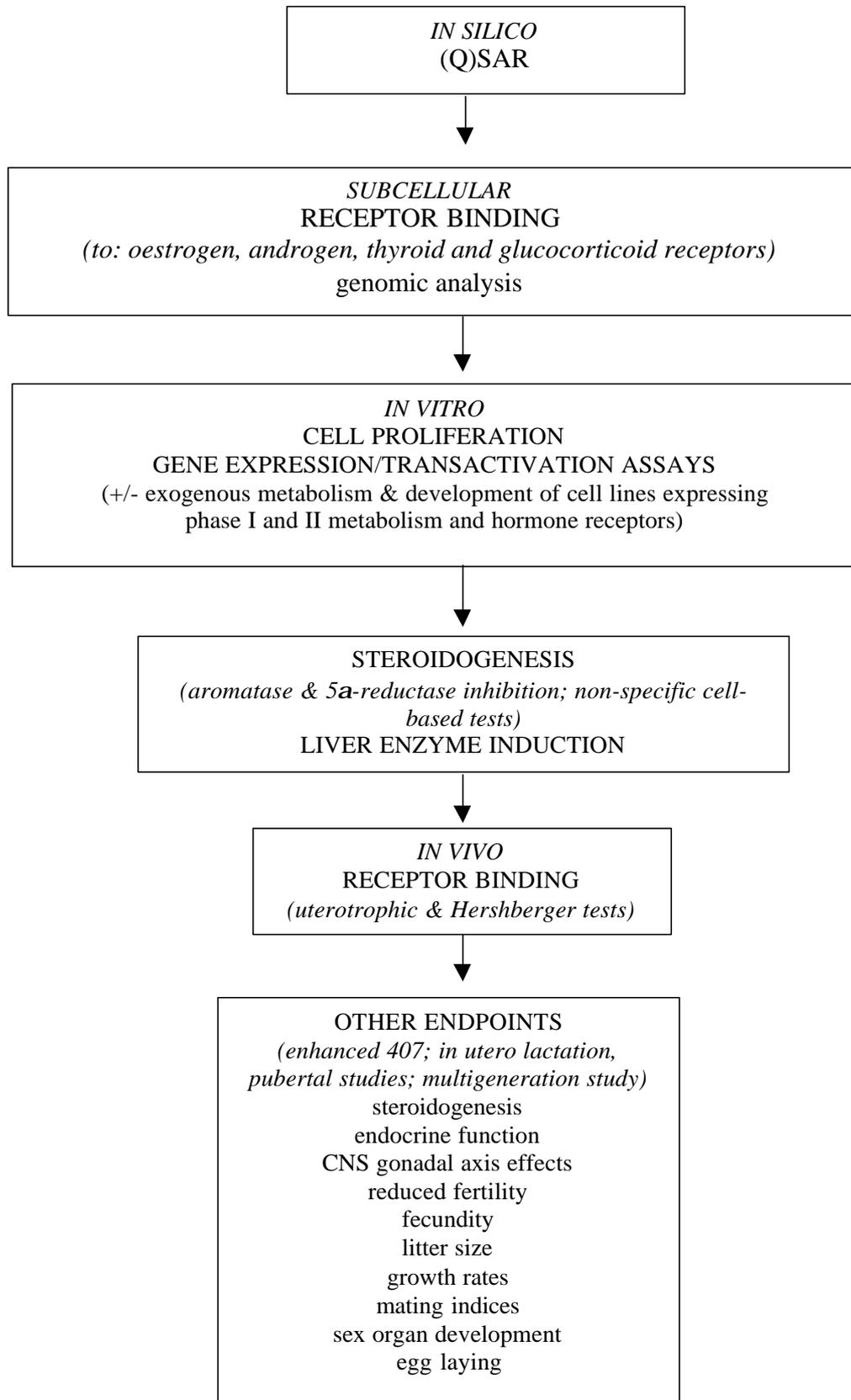sex organ development
egg laying

Table 11.1 An overview of animal assays for endocrine disruptors.

Test name	Endpoint	Mechanism of effect	Proposed role	Validation status	Comments
3-day uterotrophic assay	Change in uterine weight	Binding to the oestrogen receptor, resulting in oestrogen agonism and antagonism	Provide *in vivo* screening data for preliminary hazard identification (in EDSTAC Tier I)	OECD validation study completed; outcome awaited	Phases I and II of validation study completed and show that immature and ovariectomised animals give similar responses; few studies have included anti-oestrogens
5-day or 7-day Hershberger assay	Change in weights of seminal vesicles and ventral prostate	Binding to the androgen receptor, resulting in androgen agonism and antagonism	Provide *in vivo* screening data for preliminary hazard identification (in EDSTAC Tier I)	OECD validation study in progress	Validation study still in progress; early data suggest that no need to use castrated animals
30-day pubertal male assay with thyroid endpoints (see Stoker *et al*., 2000)	Intact 23-day-old weanling male rats are exposed to the test substance during which pubertal indices are measured. Reproductive and thyroid tissues are weighed and	Binding to various receptors associated with thyroid-related effects, steroid synthesis and 5α-reductase inhibition, to detect alterations of pubertal	Provide *in vivo* screening data for preliminary hazard identification (in EDSTAC Tier I)	None	EPA about to start a validation study. Puberty in mammalian species is a period of rapid interactive endocrine and morphological changes, and a variety of chemicals can dramatically alter pubertal development.

	evaluated histologically, and serum is taken for hormone analysis	development, thyroid function, and hypothalamic-pituitary-gonadal (HPG) system peripubertal maturation.			
Enhanced OECD TG 407 Repeat-dose 28 day study	A wide range of pathological and clinical biochemical effects, focusing on neurotoxicity, as in the standard protocol, but with extra endpoints for endocrine-sensitive effects	All possible mechanisms of ED action	None, intended to be an *in vivo* screening method to detect EDs acting by mechanisms other than by ER/AR binding	OECD validation study in progress	The relevance and reliability of the extra endpoints are not yet established, but these will be evaluated on the basis of the OECD study. Use of the assay would not increase the overall number of animals, as the protocol is part of standard testing anyway.
In utero developmental (lactation) assay	A wide range of reproductive effects in the embryo and/or offspring, following *in utero* exposure, e.g neurobehavioural endpoints, hormone levels, motor activity, malformations,	Mechanisms based on placental transfer followed by effects in the embryo and/or offspring, including the receptor-mediated effects of oestrogen,	Provide *in vivo* screening data for preliminary hazard identification (in EDSTAC Tier I)		EPA about to start a validation study.

	anogenital distance, nipple retention, time to pubity, organ weights, sperm number and motility	androgen and thyroid hormone			
Enhanced 2-generation mammalian reproductive toxicity study with extra endocrine endpoints	A variety of effects on all parameters concerned with reproduction (including fecundity, litter size, birth defects), and some endocrine organ function tests.	Agonism and antagonism to oestrogen and androgen receptors, thyroid and neuro-developmental effects.	Provide *in vivo* data for hazard identification and risk assessment purposes (in EDSTAC Tier II)		Modification of existing OECD TG 416 to include extra endpoints

Table 11.2 An overview of non-animal assays for endocrine disruptors.

Test name	Endpoint	Mechanism of effect	Proposed role	Validation status	Comments
(Q)SAR	*In silico* modelling of receptor binding	Mathematical relationship between physicochemical properties of ligand and the receptor	Compound prioritisation (HTPS)	None	Being developed by EPA laboratories, especially for pesticides; and extensively in Japan and some modelling work elsewhere
Receptor binding	Radioactive binding to hormone receptor	Specific binding of ligand to receptor	Compound prioritisation (HTPS)	None, but about to be reviewed by ICCVAM	Concerns about readiness for validation, been much activity in USA by EPA laboratories
Cell proliferation	Induction of cell division in ED-sensitive cell lines	Specific binding to receptor resulting in induction of expression of genes controlling cell division	Compound prioritisation (HTPS) EDSTAC Tier 1 test	None	Issues of specificity for EDs, repeatability, and serum binding can confound data interpretation. Phase I and II metabolism needed to be incorporated, to improve the predictivity of the assays
Reporter gene assays	Induction of transcriptional	Specific binding to receptor resulting	Compound prioritisation	None, but being peer-reviewed	The EPA has had problems with reproducibility of assays, but some

	activation of regulatory element linked to a gene whose expression is easily monitored	in induction of gene expression	(HTPS) EDSTAC Tier 1 test.	by ICCVAM	systems routinely used by industry in-house
Steroidogenesis in Leydig cells (minced testis).	Measurement of testosterone production.	Leydig cells are susceptible to androgen-binding chemicals	EDSTAC Tier I alternative/extra test.	None	A detailed review paper has been prepared for US EPA.
Aromatase assay (Brodie *et al.*, 1999)	Aromatase inhibition	Aromatase inhibitors can also inhibit oestrogen biosynthesis.	EDSTAC Tier II alternative test	None	A detailed review paper has been prepared for the EPA.

12.   SUMMARY

12.1 Introduction

This chapter brings together the summaries, conclusions and recommendations presented in this report. A distinction is made between prospects for validation on the one hand and recommendations for research and development on the other, since different organisations (and, in particular, different services of the European Commission) are responsible for funding and coordinating activities in these two areas. In addition, a number of recommendations are directed at specific organisations or expert groups, and some general recommendations are made concerning the use of alternative methods in the future REACH system.

The prospects for prevalidation and validation are organised according to whether they are considered to be short-, medium- or long-term possibilities. This reflects an optimistic assessment based on the current state-of-the-art, and does not take into account rate-limiting factors, such as the availability of funding and human resources to coordinate prevalidation and validation studies. On the basis of this assessment, an indicative time-table for the availability of validated alternative methods for chemicals testing is provided in Table 12.1.

In contrast to the time-table for validation, the recommendations for research and development are not organised into a time-frame, since it is not always possible to predict the rate at which fundamental knowledge will be acquired, or to determine when a certain test will be sufficiently well-developed for entry into the validation process.

12.2 The current situation

12.2.1 Acute lethal toxicity

Standardised basal cytotoxicity tests, such as the 3T3 NRU assay, are already widely used for non-regulatory purposes, and could be used immediately for priority setting among chemicals, and for establishing the starting dose for *in vivo* acute toxicity testing.

Future activities in this area should aim to reduce and replace the use of the rodent test for determining acute toxicity values, through the development and validation of QSARs and cytotoxicity tests.

12.2.2 Skin corrosion

Alternative methods for skin corrosion have been validated and accepted for regulatory use in the EU, so animal testing is prohibited for this endpoint. The hazard identification (classification and labelling) of skin corrosives should be based on the use of a pH test, where appropriate, and an *in vitro* test (rat skin TER assay, human skin model assay or, for qualifying test chemicals, CORROSITEX®). For risk assessment (dose-response investigations), the rat skin TER or a human skin model assay are recommended.

12.2.3 Skin irritation

In vitro methods for skin irritation testing could be used immediately for priority setting. The human skin model assays (e.g. EpiDerm™ and EPISKIN™) and the mouse SIFT appear to be the most promising of the currently-available methods.

For risk assessment purposes, there is a need to identify and evaluate the usefulness of new, mechanistically-based endpoints that are more predictive of skin irritation than are simple cytotoxicity determinations. The existing *in vitro* models also need to be improved, so that they are more representative of skin *in vivo*.

12.2.4 Eye irritation

Progress in the validation of alternative tests for eye irritation has been hampered by the lack of *in vivo* data of sufficient quality. However, data from a number of tests, including the BCOP, HET-CAM, IRE and ICE tests, are already accepted by some regulatory authorities, on a case-by-case basis, for the identification of severe eye irritants. Furthermore, the tiered testing of eye irritation is accepted by OECD Member Countries as a means of reducing and refining the use of the Draize eye test in rabbits. The use of the reference standards approach for interpreting *in vitro* data by comparison with benchmark chemicals merits futher investigation.

12.2.5 Skin sensitisation

A number of methods, including QSAR models and the DEREK skin sensitisation rulebase, reconstructed epidermis models, and dendritic cell cultures are available, and could be used for priority setting. In cases where animal testing is required, the LLNA should be used in preference to the traditional guinea-pig tests, except for those classes of chemicals for which the LLNA is not considered to be appropriate. In addition to hazard identification, the LLNA can be used for potency testing. Further fundamental work is needed before *in vitro* systems for skin sensitisation could be validated and used for regulatory purposes (classification and labelling, and dose-response assessment).

12.2.6 Respiratory sensitisation

Respiratory sensitisation is an important endpoint in the context of occupational exposure to allergenic chemicals. There is no method for respiratory sensitisation in Annex V of *Directive 67/548/EEC*, and at present no *in vitro* test is sufficiently well-developed for prevalidation. Further research and development is required to develop alternative methods in this area.

12.2.7 Biokinetics (barrier function)

Some QSARs for membrane permeability, such as those based on logP and molecular weight, may be sufficiently predictive to identify chemicals that are likely to cross biological barriers by passive diffusion. These QSARs could be used for priority setting, but chemicals that are not predicted to undergo passive diffusion should not be neglected, since non-passive modes of membrane transport are also important.

A variety of *in vitro* methods for percutaneous absorption are available and could be used for priority setting. For a complete assessment of absorption, distribution,

metabolism and excretion, absorption, *in vivo* methods can be used. For the purposes of toxicity testing, in vitro methods may be more appropriate than *in vivo* ones.

For gastrointestinal absorption, the Caco-2 culture model is considered to be sufficiently reproducible for use as a high-throughput screening system in the priority setting of chemicals. However, further work is needed to assess the predictive capacity of this system.

To assess the passage of chemicals across the BBB, a number of *in vitro* systems are under development. Further work is needed before any of these systems will be ready for prevalidation.

12.2.8 Biokinetics (xenobiotic metabolism)

A wide variety of *in vitro* tests (referred to as tier 1 tests in this report) are available for identifying metabolic pathways, metabolism-mediated toxic effects, metabolic stability and enzyme inhibition, and these could be used immediately to obtain mechanistic information. Other tests are available for assessing enzyme induction (tier 2) and polymorphic effects (tier 3). An important question for the implementation of the Chemicals Policy will be whether tier 2 and tier 3 tests are necessary, and if so, at what tonnage level of production/importation.

12.2.9 Biokinetics (physiologically-based biokinetic modelling)

PBBK models describe the processes of absorption, distribution, metabolism and excretion by integrating physicochemical, physiological and *in vitro* data. They can be used to determine target organ/system doses and to extrapolate between routes of exposure and between species. To date, most of the models have been developed for application to specific chemicals, so there is a need to develop and validated generic models, applicable to broad groups of chemicals.

12.2.10 Repeat-dose toxicity

A wide range of endpoints are investigated in *in vivo* chronic toxicity studies, so an integrated approach to chronic toxicity testing, based on the use of alternative methods with complementary endpoints, will need to be developed in order to reduce the current reliance on chronic animal tests. At present, a variety of *in vitro* systems, derived mainly from the liver, kidney and brain, are being developed. Considerable investment at the research level is needed to maintain progress in this area.

Due to the nature of long-term toxicity testing and the need to maintain *in vitro* systems with physiological characteristics similar to the *in vivo* situation, it should be recognised that *in vitro* systems for chronic toxicity will be more complicated than *in vitro* systems for acute toxicity. This should not preclude the development of long-term *in vitro* systems, and their development should not be focused on their ability to permit high-throughput screening.

12.2.11 Target organ and system toxicity

For the assessment of target organ and system toxicity, attention should be primarily focused on potential effects on the liver, the kidneys, the nervous system and the endocrine system.

Liver function can be affected by metabolism-mediated toxicity (see 11.2.8), while kidney function can also be adversely affected by loss of barrier function. Renal trans-epithelial resistance and paracellular permeability are reproducible endpoints for assessing the function of the kidney barrier.

A large number of *in vitro* systems are being developed to investigate the wide variety of endpoints associated with neurotoxicity. This is an area where a reduction in animal use is likely to result from the integrated use of alternative tests with complementary endpoints. An important question in relation to the chemicals policy is the extent to which all mechanisms need to be assessed, since, from a strategic viewpoint, it is possible that an assessment of barrier function, combined with assessments of basal cytotoxicity and energy metabolism, might be sufficient to identify substances of concern.

12.2.12 Genotoxicity and carcinogenicity

QSAR, expert system and *in vitro* approaches could be used immediately to prioritise chemicals for further testing on the basis of their potential genotoxicity and carcinogenicity. For the definitive assessment of a genotoxicity, negative results in a bacterial assay for gene mutation and in a mammalian assay for clastogenicity/aneugenicity are normally regarded as sufficient evidence for a lack of genotoxic potential

For carcinogenicity testing, the rodent bioassay is not suitable for the testing of large numbers of chemicals, not only for scientific reasons, but also because of the considerable time and cost involved, as well as the large numbers of animals required. Therefore, it is important that alternative methods for the definitive assessment of carcinogencity are further developed and then subjected to validation.

12.2.13 Reproductive toxicity

Since it is not possible to model the whole of the reproductive cycle *in vitro*, the main components need to be studied individually and then integrated in the form of a testing strategy.

For the assessment of embryotoxic potential, the EST, the WEC and the MM tests can be used immediately. However, the EST is the only one of the three which could have a relatively high throughput, and which would not involve the killing of large numbers of pregnant animals.

Considerable effort is required to develop and validate alternative tests for assessing the adverse effects of chemicals on fertility.

It will take time to develop and validate a battery of alternative tests that cover the most important aspects of the reproductive cycle, so animal tests will continue to be required, at least for certain aspects of reproductive toxicity testing. Therefore, a short-term priority should be to refine the existing animal tests for reproductive toxicity.

12.2.14 Endocrine disruption

The mechanisms whereby EDs act are complex and diverse, and this has hindered efforts to develop reliable and relevant animal and non-animal tests for their detection. The emergence of relevant and reliable animal tests has also been hindered by the fact that few, if any, of the available animal test methods have been designed with the specific purpose of detecting EDs. They are, in fact, modifications of existing animal tests. The use of animal tests for EDs poses several important welfare problems, and there are complications due to the possible existence of a low-dose effect. The validation of currently-available non-animal approaches should be made a priority, so that they can be used for compound prioritisation, and incorporated into tiered testing strategies in such a way that they reduce or replace further testing in animals. In addition, there is much research to be undertaken to assess the need for non-animal assays to cover the full spectrum of possible modes of action of EDs, and to provide definitive evidence of whether a substance does or does not interact with the hormonal systems of humans and animals.

12.3 Prospects for prevalidation and validation

12.3.1 Short-term prospects

1. The validation of basal cytotoxicity assays for predicting: a) rat oral and/or mouse i.p. LD50 values; and b) human lethal blood concentrations. Such a study has been initiated under the auspices of ICCVAM and ECVAM.

2. The validation of QSARs and/or expert system rulebases for skin corrosion, skin sensitisation and eye irritation.

3. The validation of modified test protocols for the human skin models, EPISKIN and EpiDerm, and for SIFT, to determine whether the validity of any of these methods can adequately distinguish acute skin irritants from non-irritants.

4. The prevalidation of *in vitro* tests for identifying metabolic pathways, metabolism-mediated toxic effects, metabolic stability and enzyme inhibition.

5. The prevalidation *in vitro* tests for enzyme induction, based on human hepatocyte cultures.

6. The prevalidation of re-aggregating brain cell cultures and the human neuroblastoma SHSY5Y cell line as complementary methods for assessing neurotoxicity.

7. The prevalidation of the complementary use of primary glial and neuronal cell cultures for assessing neurotoxicity.

8. The validation of (Q)SAR models and cell culture systems for predicting receptor binding of potential endocrine disrupting chemicals.

12.3.2 Medium-term prospects

1. The validation of QSAR models for barrier function (skin barrier, gastrointestinal barrier and blood-brain barrier), following an assessment of their mechanistic relevance.

2. The validation of QSAR models for skin irritation.

3. The prevalidation of *in vitro* models of the GI barrier and the BBB.

4. The validation of *in vitro* tests for identifying metabolic pathways, metabolism-mediated toxic effects, metabolic stability and enzyme inhibition.

5. The validation of *in vitro* tests for induction, based on human hepatocyte cultures.

6. The prevalidation of models for evaluating the effects of enzyme polymorphism on metabolism.

7. The validation of algorithms for predicting biokinetic processes, such as *in vivo* metabolic clearance.

8. The validation of receptor-binding assays for non-genotoxic carcinogens.

9. The validation of QSAR models and other computer-based approaches for predicting genotoxicity and carcinogenicity.

10. The validation of a modified EST, refined with target cell-specific endpoints and human embryonic stem cells.

11. The prevalidation of methods for metal-induced infertility and spermiotoxicity.

12. The prevalidation of a Leydig cell line test for assessing adverse affects on male fertility.

12.3.3 Long-term prospects

1. The validation of new *in vitro* methods for skin sensitisation.

2. The validation of new *in vitro* methods for respiratory sensitisation.

3. The validation of models for evaluating the effects of enzyme polymorphism on metabolism.

4. The validation of computer-based systems for predicting metabolism from chemical structure.

5. The prevalidation of suitably-developed models for assessing long-term toxicity.

12.4 Recommendations for research and development activities

12.4.1 Computer modelling

1. The further development of QSAR models and/or expert system rulebases for predicting *in vitro* cytotoxicity, skin corrosion, skin irritation, eye irritation, genotoxicity and carcinogenicity.

2. The further development of QSAR models and/or expert system rulebases for barrier function (skin barrier, gastrointestinal barrier, blood-brain barrier and the blood-testis barrier) and for predicting chemical effects on metabolism.

3. The investigation of computer-based methods for the clustering of chemicals according to their physicochemical characteristics, to produce an appropriate clustering method for prioritising existing substances for assessment.

4. The development of a user-friendly software package for the prediction of target organ/tissue distribution. The aim should be to make it possible for the user to enter a SMILES code or CAS number for a given chemical, and obtain a qualitative indication of the major target organs and tissues. This information could then be used to indicate the *in vitro* tests that could be performed for a chemical known to enter the body by the dermal route of exposure.

5. The further development of algorithms for predicting *in vivo* metabolic clearance from *in vitro* data.

12.4.2 Acute dermal and ocular toxicity

6. The further investigation of the applicability of the reference standards approach to the validation of *in vitro* tests for eye irritation.

7. The identification of new, mechanistically-based endpoints for skin and eye irritation, through the application of genomics and proteomics.

12.4.3 Sensitisation

8. The development of systems for skin sensitisation testing, based on an enhanced understanding of underlying biological mechanisms, including protein binding assays, human reconstructed epidermis models, and human dendritic cell cultures, which represent promising cell-based approaches.

9. The further development of *in vitro* methods for respiratory sensitisation.

12.4.4 Biokinetics (barrier function)

10. The further development of reconstituted human skin models for percutaneous absorption testing, to make their barrier properties similar to those found *in vivo*.

11. An evaluation of the feasibility of predicting bioavailability from *in vitro* data, and not just the fraction absorbed.

12. The further investigation of the expression of transport/efflux proteins in cell lines derived from the human gastrointestinal tract, and the influence of such transporters on absorption.

13. Research on the effects of anti-transport mechanisms (mdr, P-gp) and gut wall metabolism (CYP3A4) on bioavailability.

14. The further investigation of co-cultures consisting of cell lines with enterocytic markers and cell lines with mucus secretory functions, to understand the effects of mucus on the absorption rate.

15. The further optimisation of a test protocol for the blood-brain barrier that involves primary endothelial cells co-cultured with primary astrocytes.

16. The design and evaluation of a battery of *in vitro* assays for predicting the distribution of compounds to the brain. Such a battery could include, for example, measurements of protein binding and clearance, and the use of MDCK cells transfected with transporter proteins (such as mdr-1).

17. The improvement and/or establishment of new cell models of the blood-brain barrier that have characteristics more consistent with the *in vivo* situation, with an emphasis on the use of human cell lines.

12.4.5 Biokinetics (xenobiotic metabolism)

18. The further improvement of techniques for the cryopreservation of human hepatocytes.

19. The further development of *in vitro* models for evaluating the effects of enzyme polymorphism on metabolism.

12.4.6 Target organ and system toxicity

20. The identification of relevant biomarkers of exposure and effect for target organ/system toxicity testing, The endpoints selected should cover general cytotoxic mechanisms and cell-type-specific mechanisms of toxicity. There is a need to develop non-invasive methods for determining such endpoints.

21. An assessment of the immunological basis of target organ toxicity.

22. The further investigations of the use of tissue renal slices as a model for studying the mechanisms of nephrotoxicity.

23. The immortalisation of primary proximal tubular cell cultures, and further characterisation of the new cell lines, to establish whether they retain the characteristics of their *in vivo* precursor cells through several passages.

24. The further development of the use of HK-2 cells (human proximal tubular epithelial cell line) and LLC-PK1 cells (porcine proximal tubular epithelial cell

line), grown under long-term conditions, and an assessment of their usefulness for transport across the kidney barrier.

25. An assessment of the extent to which loss of barrier function can account for kidney damage.

26. An evaluation of the usefulness of genetically-modified PC12 cell lines as a component in a testing strategy for neurotoxicity.

27. The further development of *in vitro* models for evaluating the mechanisms of neurotoxicity, such as genetically-engineered cell lines, and re-aggregating cultures of human embryonic stem cell lines.

28. The evaluation of genomics, proteomics and new electrophysiological and biochemical profiling methods for neurotoxicity testing

29. The development and evaluation of a tiered testing strategy for neurotoxicity, taking into account the results obtained in previous studies.

30. The development of systems for predicting toxicotolerance.

31. The identification of reference compounds suitable for the development of relevant and reliable *in vitro* procedures for long-term toxicity testing.

32. The development of non-invasive imaging techniques for detecting long-term effects *in vivo*.

33. The development of methods for long-term toxicity testing, employing human-based hepatic, renal and neuronal cell lines expressing a wide range of drug and xenobiotic metabolising enzymes and transport molecules, as an alternative to primary cultures.

34. The further development of long-term culture methods (several weeks to months) for hepatic and renal epithelial and endothelial cells, as well as for neurons and glial cells.

35. The further development of co-culture systems for long-term toxcity testing: a) neurons with glia cells; b) hepatocytes with monocytic, Ito cells and/or endothelial cells; c) renal epithelial cells with renal microvascular endothelial cells; and d) renal glomerular mesangial cells with glomerular endothelial and/or glomerular epithelial cells.

36. The refinement of perfusion culture systems for chronic toxicity testing, with an emphasis on miniaturisation and practicability, to provide effective, technically simple and sensitive systems for assessing the effects of test compounds with biologically relevant endpoints.

37. The further development of metabolically competent, genetically-engineered cell lines, grown on microporous supports and continuously perfused with

conventional culture medium, as models for evaluating the effects of continuous low doses and long-term exposure in the liver, kidney and neuronal tissue.

38. The further development of the available human renal co-culture systems (e.g. MDCK cells with metabolically-competent cells), since epithelial and endothelial cells can be successfully grown in combination under static culture conditions and under continuous medium perfusion.

12.4.7 Genotoxicity and carcinogenicity

39. A review of the current validation and regulatory status of aneuploidy and micronucleus assays in mammalian cells, to include an assessment of the range of aneugens detected by both methods, with a view to eliminating redundancy.

40. A review of the current validation and regulatory status of rodent cell transformation assays (especially the SHE cell and Balb C systems).

41. A review of the need to conduct *in vivo* genotoxicity testing, especially for chemicals that are not intended for human consumption or direct exposure, and those for which indirect exposure is expected to be negligible.

42. The development of receptor-binding assays for non-genotoxic carcinogens.

43. The development of human cell-based cell transformation assays for carcinogenicity testing.

44. Further research concerning: a) the known modes of action of non-genotoxic carcinogens; b) their relevance to human hazard; and c) the available and required non-animal testing strategies for detecting and characterising those carcinogens that are relevant.

12.4.8 Reproductive toxicity

45. The identification of the most predictive toxicological endpoints for use in fertility testing with semen analysis.

46. An evaluation of the use of Sertoli cell lines and Sertoli cell co-cultures for assessing adverse effects on male fertility.

47. The development of a testing strategy for developmental toxicity, which covers not only malformations, but also other manifestations of developmental toxicity, such as growth retardation and embryolethality.

48. An evaluation of the use of primary cultures of spermatogonia and oocytes for the assessment of adverse effects on fertility.

49. The further development of granulosa cells and theca cells as toxicological screening systems.

50. Research on the use of PGCs and/or PGC lines for identifying germ cell mutagens.

51. The development of a testing strategy covering all essential aspects of the male and female reproductive cycles.

12.4.9 Endocrine disruption

52. Investigations on the use of microarrays for screening for receptor binding and gene induction, coupled with the search for further relevant receptors.

53. Research on the use of *in vitro* receptor binding assays in conjunction with biokinetic modelling.

54. The assessment of the potential use of biomarkers of exposure and effect for EDs.

55. Investigations on the suitability of *in vitro* systems for yielding consistent data regarding the potential interactive effects of chemical mixtures of EDs.

56. Further research on the basic mechanisms of ED action, especially via non-receptor pathways, and the development of *in vitro* models for such processes.

57. The development of appropriate *in vitro* methods for measuring steroidogenesis in males and females and for investigating endocrine function.

58. The development of approaches for detecting ED activity by using exogenous metabolising systems and cell culture systems, comprising transgenic cell lines with relevant hormone-receptor response elements, reporter gene sequences, and genes for phase I and phase II metabolism.

59. The development and application of immortalised fetal and pre-pubertal cell lines.

60. The development of an integrated testing strategy for EDs, based on the maximum use of non-animal approaches.

12.4.10 General

61. An assessment of the usefuless of genomics and protemics in toxicity testing.

62. An assessment of the replacement, reduction and refinement possibilities in relation to the REACH system.

63. The development of a strategy for assessing ecotoxicological endpoints, based on the use of alternative methods.

12.5 Recommendations to specific organisations

1. National regulatory authorities should consideration harmonising their positions on the acceptance of the BCOP, HET-CAM, IRE, ICE, and other non-animal tests for eye irritation.

2. The DEREK User Group is encouraged to make more information publicly available on the DEREK rulebase for skin sensitisation.

3. The ESAC should consider making an endorsement of a statement on the applicability of *in vitro* methods for percutaneous absorption, assessed on the basis of a weight-of-evidence approach.

4. ECVAM should organise workshops on:

a) non-genotoxic carcinogens, in conjunction with the FRAME Toxicity Committee and the International Agency for Research on Cancer, to discuss and make recommendations for: a) the known modes of action of non-genotoxic carcinogens; b) their relevance to human hazard; and c) the available and required non-animal testing strategies for detecting and characterising those carcinogens that are relevant;

b) the use of genetically-engineered cell lines for predicting metabolism-mediated genotoxicity;

c) embryotoxicity, to define the areas of application of the scientifically-validated EST; and

d) the use of hormone-producing cells for predicting the adverse effects of chemicals on fertility.

e) the use of *in vitro* systems for screening for EDs, and to define the roles of such tests for this purpose. The workshop should include experts on endocrine disruption, *in vitro* toxicologists, clinical endocrinologists, ecologists, chemists, experts in (Q)SAR, and regulators. The workshop should focus on the limitations of *in vitro* approaches, and the need for them in relation to *in vivo* tests, bearing in mind that so far, no uniquely *in vivo* positive chemicals have been identified.

5. The appropriate regulatory authorities should consider the scientific merits of a test battery for EDs, either to supplement the existing reproductive toxicity test package, or to be applied as a separate set of tests specifically for ED activity.

12.6 General recommendations for the use of alternatives in the REACH system

1. To maximise the number of alternative tests available and suitable for use in the REACH system, emphasis should be placed on the development and validation of alternative tests that are already considered to be promising, since validated and accepted tests are likely to be required by 2008, in order to meet the proposed deadline of 2012 for the testing of the 20,000 existing chemicals produced in the range of 1-10 t.p.a.

2. For the assessment of endpoints where replacement alternatives are unlikely to be available in the near future (e.g. chronic toxicity testing), attempts should be made to reduce and refine the currently required animal tests until suitable alternative methods, or batteries of such methods, have been developed and validated. To make progress in this respect, a comprehensive review should be undertaken to assess the relevance and reliability of all the animal tests that will eventually be required for chemicals testing under the REACH system, and the outcome should be published in the peer-review literature. This should also include an assessment of refinement and reduction possibilites in relation to the new legislation.

3. Testing strategies for the assessment of toxicological endpoints should be designed with respect to the known uses of chemicals and the exposure to them, at the beginning of the strategy.

4. Testing strategies should be based on a consideration of all of the Three Rs, including the use of alternative methods and the use of existing information, wherever possible.

5. Human volunteers should be used only to confirm safety. In particular, biokinetic studies on absorption, distribution, metabolism and excretion, and assessments of corrosivity, sensitisation and mutagenicity, should be undertaken prior to any human studies.

Table 12.1. An indicative time-table for the prevalidation and validation of alternative tests for chemicals testing.

Effect	EU test methods	Alternative methods	Status	Prospects for completion of validation
Acute lethal toxicity	B.1bis B.1tris B.2 B.3	1) Basal cytotoxicity tests, including the neutral red uptake (NRU) assay in BALB/c 3T3 cells and the Normal Human Keratinocytes 2) QSARs for basal cytotoxicity	1) ICCVAM-ECVAM VS in progress 2) R&D	1) Short-term: expected completion of VS by June 2003 2) Medium-term
Skin corrosion	B.40	1) Rat skin TER 2) human skin models (EPISKIN and EpiDerm) 3) CORROSITEX 4) QSARs and expert system rulebases	1-2) Validated by ECVAM and accepted at the EU level 3) Validated by ICCVAM and endorsed by ECVAM/ESAC 4) R&D	1-3) No further validation of these models is required; however, new human skin models could be validated 4) Short-term
Skin irritation	B.4	1) Human skin models (e.g EPISKIN and EpiDerm) 2) Skin Integrity Function Test (SIFT) 3) Pig ear test 4) QSARs and expert system rulebases	1-3) ECVAM PVS completed.	1-2) Short-term 3) Medium-term 4) Medium-term
Eye irritation	B.5	1) BCOP test 2) HET-CAM 3) Isolated rabbit eye (IRE) test 4) Isolated chicken eye (ICE) test 5) Fluorescein leakage assay 6) NRU assay 7) Neutral red release (NRR) assay 8) Red blood cell (RBC) haemolysis test 9) EpiOcular assay 10) EYTEX	Several VSs have been conducted, but no method is regarded as validated. Data obtained with some of these tests are accepted by regulatory authorities on a case-by-case basis	1-9) Medium-term: provided that several problems are overcome, including provision of high-quality *in vivo* data, and further definition of the reference standards approach 10) Not recommended

		11) QSARs and expert system rulebases		11) Short-term
Skin sensitisation	B.6	1) Protein-binding assays 2) Numerous cell-based systems, including dendritic cells, Langerhans cells and human skin models 3) QSARs and expert system rulebases	1-3) R&D	1-2) Medium-term 3) Short-term
Respiratory sensitisation	No Annex V method	No well-characterised system available	R&D	Long-term
Biokinetics	B.36	Numerous alternative methods are available, including: 1) *in vitro* and QSAR models for barrier function 2) *in vitro* models for metabolism 3) QSARs and expert system rulebases for metabolism 4) mathematical models for biokinetic processes	R&D Data obtained with *in vitro* models of the skin barrier are accepted by regulatory authorities on a case-by-case basis	1) Medium-term for QSAR models; long-term for *in vitro* models 2) Long-term 3-4) Long-term
Target organ/system toxicity	B.7 B.8 B.9	Numerous *in vitro* systems are being developed for cell-specific toxicity, especially for: 1) hepatotoxicity 2) nephrotoxicity 3) neurotoxicity 4) endocrine system toxicity	R&D	1) Medium-term 2-4) Long-term
Chronic toxicity	B.26 B.27 B.28 B.29 B.30	Numerous *in vitro* systems are being developed, especially for: 1) hepatotoxicity 2) nephrotoxicity 3) neurotoxicity	R&D	1-3) Long-term
Genotoxicity	B.10-B25, B.39	Tests for gene mutation, including: 1) bacterial reverse mutation assay (Ames) 2) mammalian cell mutation assay (e.g.	Data obtained with tests 1-4 are accepted by regulatory authorities on a case-by-case	

		mouse lymphoma assay)	basis	
		Tests for clastogenicity, including: 3) *In vitro* cytogenetic assay using metaphase analysis 4) *In vitro* micronucleus assay	4) ongoing validation	
		5) QSARs and expert system rulebases	5) R&D	5) Medium-term
Carcinogenicity	B.21 B.32 B.33	1) Syrian Hamster embryo (SHE) cell transformation assay 2) Balb C/3T3 cell transformation assay 3) 3D invasiveness assays	1-3) R&D	1-3) Long-term
Reproductive toxicity	B.31 B.34 B.35	1) Whole-embryo culture (WEC) assay 2) Embryonic stem cell test (EST) 3) Micromass (MM) assay 4) Sertoli cell lines 5) Leydig cell lines 6) Sperm motility and morphology test 7) FETAX	1-3) Validated by ECVAM	1-3) Ready for regulatory consideration; further assessment of these models should focus on improvement of their prediction models 4-6) Long-term 7) Not recommended

Footnotes to Table 12.1

The EU Test Methods are described in Table 1.4.
Short-term = by end of 2003; Medium-term = by end of 2006; Long-term = by end of 2010.
PVS = prevalidation study; R&D = research and development; VS = validation study.

The time-frame for the short-, medium- and long-term prospects represents an optimistic assessment based on the current state-of-the-art. In practice, the achievement of progress in line with the time-frame will depend on the availability of sufficient human and financial resources, and on the coordination at the EU level of complementary activities. The realisation of the medium-term and long-term prospects will depend on adequate test development in the short-term and medium-term. The time-frame takes into account the need to perform prevalidation prior to formal validation, where appropriate.

13.   REFERENCES

Chapter 1. Background

Anon (1967). Council Directive *67/548/EEC* of 27 June 1967 on the approximation of laws, regulations and administrative provisions relating to the classification, packaging and labelling of dangerous substances. *Official Journal of the European Communities* **196**, 16.8.1967.

Anon (1976). Council Directive *76/769/EEC* of 27 July 1976 on the approximation of the laws, regulations and administrative provisions of the Member States relating to restrictions on the marketing and use of certain dangerous substances and preparations. *Official Journal of the European Communities* **L262**, 27.9.1976.

Anon (1988). Council Directive *88/379/EEC* of 7 June 1988 on the approximation of laws, regulations and administrative provisions of the Member States relating to the classification, packaging and labelling of dangerous preparations. *Official Journal of the European Communities* **L187**, 16.7.1988.

Anon (1986). Council Directive *86/609/EEC* of 24 November 1986 on the approximation of laws, regulations and administrative provisions of the Member States regarding the protection of animals used for experimental and other scientific purposes. *Official Journal of the European Community* **L358**, 18.12.1986.

Anon (1991). *Establishment of a European Centre for the Validation of Alternative Methods (ECVAM).* Communication from the Commission to the Council and the European Parliament. Brussels: Commission of the European Communities.

Anon (1993). Council Regulation (EEC) No 793/93 of 23 March 1993 on the evaluation and control of risks of existing substances. *Official Journal of the European Communities* **L84**, 5.4.1993.

Anon (2001). White Paper on a Strategy for a Future Chemicals Policy. Brussels: Commission of the European Communities. Available on the internet: http://europa.eu.int/comm/environment/chemicals/whitepaper.htm.

Anon (2001a). Report of the 2355th session of the Council "Environment", held on 7 June 2001. pp. 10-16. Luxembourg: Council of Ministers. Available on the internet: http://ue.eu.int/newsroom/

Anon (2001b). Report on the Commission White Paper on Strategy for a future Chemicals Policy. Document No. A5-0356/2001. Adopted in plenary session on 15 November, 2001. Brussels: European Parliament.

Anon (2001c). Opinion of the Economic and Social Committee on the White Paper - Strategy for a future Chemicals Policy (COM(2001) 88 final). Document No. NAT/117. Adopted in plenary session on 17 October, 2001. Brussels: Economic and Social Committee.

BUAV (2001). *Action to End Animal Testing. The Way Forward*. 19 pp. London: British Union for the Abolition of Vivisection.

Festing, M.F.W., Baumans, V., Combes, R.D., Halder, M., Hendriksen, C.F.M., Howard, B.R., Lovell, D.P., Moore, G.J., Overend, P. & Wilson, M.S. (1998). Reducing the use of laboratory animals in biomedical research: problems and possible solutions. *ATLA* **26**, 283-301.

Green, S., Goldberg, A.M. & Zurlo, J. (2001). The TestSmart HPV program – development of an integrated approach for testing high production volume chemicals. *Regulatory Toxicology and Pharmacology* **33**, 105-109.

Hendriksen, C.F.M & Morton, D.B. (eds) (1999). *Humane Endpoints in Animal Experiments for Biomedical Research*. London: The Royal Society of Medicine press.

IEH (2001a). *Testing Requirements for Proposals under the EC White Paper 'Strategy for a Future Chemicals Policy'*. 25 pp. Leicester, UK: Medical Research Council Institute for Environment and Health. Available on the internet: http://www.le.ac.uk/ieh/pdf/w6.pdf

IEH (2001b). *Assessment of the Feasibility of Replacing Current Regulatory In Vivo Toxicity Tests with In Vitro Tests within the Framework Specified in the EU White Paper 'Strategy for a Future Chemicals Policy'*. 40 pp. Leicester, UK: Medical Research Council Institute for Environment and Health. Available on the internet: http://www.le.ac.uk/ieh/pdf/w10.pdf

Russell, W.M.S. & Burch, R.L. (1959). *The Principles of Humane Experimental Technique*. 238 pp. London: Methuen.

Smyth, D.H. (1978). Alternatives to animal experiments. London: Scolar Press-Royal Defence Society, 1978.

Worth, A.P. & Balls, M. (2001). The role of ECVAM in promoting the regulatory acceptance of alternative methods in the European Union. *ATLA* **29**, 525-535.

Chapter 2. Principles and procedures of validation

Archer, G. (1996). No statistics, no science - let alone validation studies. *ATLA* **24**, 453-456.

Archer, G., Balls, M., Bruner, L.H., Curren, R.D., Fentem, J.H., Holzhütter, H-G., Liebsch, M., Lovell, D.P. & Southee, J.A. (1997). The validation of toxicological prediction models. *ATLA* **25**, 505-516.

Balls, M. (1997). Defined structural and performance criteria would facilitate the validation and acceptance of alternative test procedures. *ATLA* **25**, 483-484.

Balls, M., Blaauboer, B., Brusick D., Frazier J., Lamb, D., Pemberton, M., Reinhardt, C., Roberfroid, M., Rosenkranz, H., Schmid, B., Spielmann, H., Stammati, A-L. &

Walum, E. (1990). Report and recommendations of the CAAT/ERGATT workshop on the validation of toxicity test procedures. *ATLA* **18**, 313-337.

Balls, M., Blaauboer, B.J., Fentem, J.H., Bruner, L., Combes, R.D., Ekwall, B., Fielder, R.J., Guillouzo, A., Lewis, R.W., Lovell, D.P., Reinhardt, C.A., Repetto, G., Sladowski, D., Spielmann, H. & Zucco, F. (1995a). Practical aspects of the validation of toxicity test procedures. The report and recommendations of ECVAM workshop 5. *ATLA* **23**, 129-147.

Balls, M., Botham, P.A., Bruner, L.H. & Spielmann, H. (1995b). The EC/HO international validation study on alternatives to the Draize eye irritation test. *Toxicology in Vitro* **9**, 871-929.

Balls, M. & Fentem, J.H. (1997). Progress toward the validation of alternative tests. *ATLA* **25**, 33-43.

Balls, M. & Fentem, J.H. (1999). The validation and acceptance of alternatives to animal testing. *Toxicology in Vitro* **13**, 837-846.

Balls, M. & Karcher, W. (1995). The validation of alternative test methods. *ATLA* **23**, 884-886.

Barratt, M.D., Brantom, P.G., Fentem, J.H., Gerner, I., Walker, A.P. & Worth, A.P. (1998). The ECVAM international validation study on *in vitro* tests for skin corrosivity. 1. Selection and distribution of the test chemicals. *Toxicology in Vitro* **12**, 471-482.

Botham, P.A., Chamberlain, M., Barratt, M.D., Curren, R.D., Esdaile, D.J., Gardner, J.R., Gordon, V.C., Hildebrand, B., Lewis, R.W., Liebsch, M., Logemann, P., Osborne, R., Ponec, M., Régnier, J.F., Steiling, W., Walker, A.P. & Balls, M. (1995). A prevalidation study on *in vitro* skin corrosivity testing. The report and recommendations of ECVAM workshop 6. *ATLA* **23**, 219-255.

Bruner, L.H., Carr, G.J., Chamberlain, M. & Curren, R.D. (1996a). Validation of alternative methods for toxicity testing. *Toxicology in Vitro* **10**, 479-501.

Bruner, L., Carr, G., Chamberlain, M. & Curren, R. (1996b). No prediction model – no validation study. *ATLA* **24**, 211-217.

Cooper-Hannan, R., Harbell, J.W., Coecke, S., Balls, M., Bowe, G., Cervinka, M., Clothier, R., Hermann, F., Klahm, L.K., de Lange, J., Liebsch, M. & Vanparys, P. (1999). The principles of good laboratory practice: application to *in vitro* toxicology studies. The report and recommendations of ECVAM workshop 37. *ATLA* **27**, 539-577.

Curren, R.D., Southee, J.A., Spielmann, H., Liebsch, M., Fentem, J.H. & Balls, M. (1995). The role of prevalidation in the development, validation and acceptance of alternative methods. *ATLA* **23**, 211-217.

ECVAM (1998a). Statement on the scientific validity of the EPISKIN[TM] test (an *in vitro* test for skin corrosivity). *ATLA* **26**, 277-280.

ECVAM (1998b). Statement on the scientific validity of the rat skin transcutaneous electrical resistance (TER) test (an *in vitro* test for skin corrosivity). *ATLA* **26**, 275-277.

ECVAM (2001). Statement on the application of the EPIDERM[®] human skin model for skin corrosivity testing. *ATLA* **28**, 365-367.

Fentem, J.H., Archer, G.E.B., Balls, M., Botham, P.A., Curren, R.D., Earl, L.K., Esdaile, D.J., Holzhütter, H.G. & Liebsch, M. (1998). The ECVAM international validation study on *in vitro* tests for skin corrosivity. 2. Results and evaluation by the management team. *Toxicology in Vitro* **12**, 483-524.

Fentem, J.H. & Balls, M. (1997). The ECVAM approach to validation. In *Animal Alternatives, Welfare and Ethics. Proceedings of the 2nd World Congress on Alternatives and Animal Use in the Life Sciences, Utrecht, The Netherlands, 20-24 October 1996* - Developments in Animal and Veterinary Sciences, Volume 27 (ed. L.F.M. van Zutphen & M. Balls), pp. 1083-1089. Elsevier, Amsterdam, The Netherlands.

Fentem, J. H., Briggs, D., Chesné, C., Elliott, G. R., Harbell, J. W., Heylings, J. R., Portes, P., Roguet, R., van de Sandt, J. J. M. & Botham, P. A. (2001). A prevalidation study on *in vitro* tests for acute skin irritation: results and evaluation by the Management Team. *Toxicology in Vitro* **15**, 57-93.

Genschow, E., Scholz, G., Brown, N.A., Piersma, A.H., Brady, M., Clemann, N., Huuskonen, H., Paillard, F., Bremer, S. & Spielmann, H. (2000). Development of prediction models for three *in vitro* embryotoxicity tests in an ECVAM Validation Study. *In Vitro & Molecular Toxicology* **13**, 51-65.

Holzhütter, H-G., Archer, G., Dami, N., Lovell, D.P., Saltelli, A. & Sjöström, M. (1996). Recommendations for the application of biostatistical methods during the development and validation of alternative toxicological methods. ECVAM Biostatistics Task Force Report 1. *ATLA* **24**, 511-530.

Holzhütter, H-G. & Spielmann, H. (1996). The use of prediction models in toxicology: what should they predict ? *ATLA* **24**, 881-884.

Liebsch, M., Traue, D., Barrabas, C., Spielmann, H., Uphill, P., Wilksins, S., McPherson, J.P., Wiemann, C., Kaufmann, Remmele, M. & Holzhütter, H-G. (2000). The ECVAM prevalidation study on the use of EpiDerm for skin corrosivity testing. *ATLA* **28**, 371-401.

NIH (2001a). Report of the international workshop on *in vitro* methods for assessing acute systemic toxicity. NIH Publication 01-4499. Research Triangle Park, NC, USA: NIEHS.

NIH (2001b). Guidance document on using *in vitro* data to estimate *in vivo* starting doses for acute toxicity. NIH publication 01-4500. Research Triangle Park, NC, USA: NIEHS.

NIH (1997). Validation and Regulatory Acceptance of Toxicological Test Methods. A Report of the *ad hoc* Interagency Coordinating Committee on the Validation of Alternative Methods. NIH Publication 97-3981. 105 pp. Research Triangle Park, NC, USA: NIEHS.

NIH (1999). Evaluation of the Validation Status of Toxicological Methods: General Guidelines for Submissions to ICCVAM (Revised, October 1999). NIH Publication 99-4496. 44 pp. Research Triangle Park, NC, USA: NIEHS.

OECD (1996). *Final Report of the OECD Workshop on Harmonization of Validation and Acceptance Criteria for Alternative Toxicological Test Methods*. 60 pp. Paris, France: OECD.

Russell, W.M.S. & Burch, R.L. (1959). *The Principles of Humane Experimental Technique*. Methuen, London.

Spielmann, H., Balls, M., Dupuis, J., Pape, W.J., Pechovitch, G., de Silva, O., Holzhütter, H.G., Clothier, R., Desolle, P., Gerberick, F., Liebsch, M., Lovell, W.W., Maurer, T., Pfannenbecker, U., Potthast, J.M., Csato, M., Sladowski, D., Steiling, W., Brantom, P. (1998). The International EU/COLIPA In Vitro Phototoxicity Validation Study: Results of Phase II (Blind Trial). Part 1: The 3T3 NRU Phototoxicity Test. *Toxicology In Vitro* **12**, 305-327.

Spielmann, H., Liebsch, M., Kalweit, S., Moldenhauer, F., Wirnsberger, T., Holzhütter, H-G., Schneider, B., Glaser, S., Gerner, I., Pape, W.J.W., Kreiling, R., Krauser, K. & Miltenburger, H.G., Steiling, W., Luepke, N.P., Muller, N., Kreuzer, H., Murmann, P., Spengler, J., Bertram-Neis, E., Siegemund, B. & Wiebel, F.J. (1996). Results of a validation study in Germany on two *in vitro* alternatives to the Draize eye irritation test, the HET-CAM test and the 3T3-NRU cytotoxicity test. *ATLA* **24**, 741-858.

Worth, A.P. & Balls, M. (2001a). The importance of the prediction model in the development and validation of alternative tests. *ATLA* **29**, 135-143.

Worth, A.P. & Balls, M. (2001b). The role of ECVAM in promoting the regulatory acceptance of alternative methods in the European Union. *ATLA* **29**, 525-535.

Worth, A.P. & Cronin, M.T.D. (2001). Prediction models for eye irritation potential based on endpoints of the HETCAM and neutral red assays. *In Vitro & Molecular Toxicology* **14**, 143-156.

Chapter 3. The scientific basis of chemical risk assessment

Balls, M. & Fentem, J.H. (1992). The use of basal cytotoxicity and target organ toxicity tests in hazard identification and risk assessment. *ATLA* **20**, 368-388.

Barratt, M.D. (2000). Principles of toxicity prediction from chemical structure. In *Progress in Reduction, Refinement and Replacement in Animal Experimentation,* (ed. M. Balls, A-M. van Zeller & M.E. Halder). pp. 449-456. Amsterdam: Elsevier Science, BV.

Blaauboer, B.J., Barratt, M.D. & Houston, J.B. (1999). The Integrated Use of Alternative Methods in Toxicological Risk Evaluation. ECVAM Integrated Testing Strategies Task Force Report 1. *ATLA* **27**, 229-237.

Bremer, S., Worth, A.P., Paparella, M., Bigot, K., Kolossov, E., Fleischmann, B.K., Hescheler, J. & Balls, M. (2001). Establishment of an *in vitro* reporter gene assay for developmental cardiac toxicity. *Toxicology in Vitro* **15**, 215-223.

DeJongh, J., Verhaar, H.J. & Hermens, J.L. (1997). A quantitative property-property relationship (QPPR) approach to estimate *in vitro* tissue-blood partition coefficients of organic chemicals in rats and humans. *Archives of Toxicology* **72**, 17-25.

DeJongh, J., Verhaar, H.J.M., Hermens, J.L.M. (1998) Role of kinetics in acute lethality of nonreactive volatile organic compounds (VOCs). *Toxicological Science* **45**, 26-32.

Fentem, J. H., Archer, G. E. B., Balls, M., Botham, P. A., Curren, R. D., Earl, L. K., Esdaile, D. J., Holzhütter, H. G. & Liebsch, M. (1998). The ECVAM international validation study on *in vitro* tests for skin corrosivity. 2. Results and evaluation by the Management Team. *Toxicology in Vitro* **12**, 483-524.

EPA (1999). *Guidelines for Carcinogen Risk Assessment*. Risk Assessment Forum. NCEA-F-0644. Washington, DC, USA: Environmental Protection Agency.

Frazier, J.M. (1994). The role of mechanistic toxicology in test method validation. *Toxicology in Vitro* **8**, 787-791.

Hanway, R.H. & Evans, P.F. (2000). Read-across of toxicological data in the notification of New Chemicals. *Toxicology Letters* **116** (S1), 61.

IPCS (1978). *Principles and methods for evaluating the toxicity of chemicals.* Part I. International Programme on Chemical Safety. Environmental Health Criteria 6. Geneva: WHO.

Kroes, R. & Feron, V.J. (1990). Toxicity testing: strategies and conduct. In *Progress in Predictive Toxicology* (ed. D.B. Clayson, I.C. Munro, P. Shubik & J.A. Swenberg). pp 15-39. New York: Elsevier Science.

OECD (1998). *Harmonized Integrated Hazard Classification System for Human Health and Environmental Effects of Chemical Substances.* Organisation for Economic Cooperation and Development, Paris.

Soffers, A.E.M.F., Boersma, M.G., Vaes, W.H.J., Vervoort, J., Tyrakowska, B., Hermens, J.L.M. & Rietjens, I.M.C.M. (2001). Computer-modeling-based QSARs for

analyzing experimental data on biotransformation and toxicity. *Toxicology in Vitro* **15**, 539 –551.

Spielmann, H., Genschow, E., Liebsch, M. & Halle, W. (1999). Determination of the starting dose for acute oral toxicity (LD50) testing in the up and down procedure (UDP) from cytotoxicity data. *ATLA* **27**, 957-966.

Stingele, S., Coecke, S., Nicotera, P. & Balls, M. (1999). EC Patent filed 27th of August 1999 on *Genetically engineered cell lines, and their uses, in particular for neurotoxicity testing*. Patent No. 99402139.2-2106/International N0 PCT/EP00/08223. Rijwijk, The Netherlands: European Patent Office.

Thomas, R.S., Rank, D.R., Penn, S.G., Zastrow, G.M., Hayes, K.R., Pande, K., Glover, E., Silander, T., Craven, M.W., Reddy, J.K., Jovanovich, S.B. & Bradfield, C.A. (2001). Identification of toxicologically predictive gene sets using cDNA microarrays. *Molecular Pharmacology* **60**, 1189-1194.

Chapter 4. Acute lethal toxicity

Balls, M. & Clothier, R.H. (1992a). Cytotoxicity assays for intrinsic toxicity and irritancy. In *In Vitro Methods for Toxicology* (ed. R.R. Watson). pp 37-52. Boca Taton, FL: CRC Press.

Balls, M. & Fentem, J.H. (1992b). The use of basal cytotoxicity and target organ toxicity tests in hazard identification and risk assessment. *ATLA* **20**, 368-388.

Balls, M., Botham, P.A., Bruner, L.H. & Spielmann, H. (1995). The EC/HO international validation study on alternatives to the Draize eye irritation test for classification and labelling of chemicals. *Toxicology in Vitro* **9**, 871-929.

Brantom, P. G., Bruner, L.H., Chamberlain, M., DeSilva, O., Dupuis, J., Earl, L.K., Lovell, D.P., Pape, W.J.W., Uttley, M., Bagley, D.M., Baker, F.W., Baker, F.W., Bracher, M., Courtellemont, P., Declercq, I., Freeman, S., Steiling, W., Walker, A.P., Carr, G.J., Dami, N., Thomas, G., Harbell, J., Jones, P.A., Pfannenbecker, U., Southee, J.A., Tcheng, M., Argembeaux, H., Castelli, D., Clothier, R.H., Esdaile, D.J., Itigaki, H., Jung, K., Kasai, Y., Kojima, H., Kristen, U., Larnicol, M., Lewis, R.W., Marenus, K., Moreno, O., Peterson, A., Rasmussen, E.S., Robles, C. & Stern, M. (1997). A summary report of the COLIPA international validation study on alternatives to the Draize rabbit irritation test. *Toxicology in Vitro* **11**, 141-179.

Calleja, M.C., Geladi, P. & Persoone, G. (1994). QSAR models for predicting the acute toxicity of selected organic chemicals with diverse structures to aquatic non-vertebrates and humans. S*AR & QSAR in Environmental Research* **2**,193-234.

Clemedson, C. & Ekwall, B. (1999). Overview of the final MEIC results: I. The *in vitro-in vivo* evaluation. *Toxicology in Vitro* **13**, 1-7.

Clemedson, C. *et al*. (2000). MEIC evaluation of acute systemic toxicity. Part VII. Prediction of human toxicity by results from testing of the first 30 reference chemicals with 27 further *in vitro* assays. *ATLA* **28**, Supplement 1, 159-200.

Clothier, R.H., Hulme, L.M., Sith, M. & Balls, M. (1987). Comparison of the *in vitro* cytotoxicities and acute in vivo toxicities of 59 chemiclas. *Molecular Toxicology* **1**, 571 -577.

Clothier, R.H., Hulme, A.B., Ahmed, H.L., Reeves, M. & Balls, M. (1988). *In vitro* cytotoxicity of 150 chemicals to 3T3-L1 cells, assessed by the FRAME Kenacid Blue Method. *ATLA* **16**, 84-95.

Clothier, R.H., Willshaw, A., Cox, H., Garle, M., Bowler, H. & Combes, R. (1999). The use of human keratinocytes in the EU/COLIPA international *in vitro* phototoxicity test validation study and the ECVAM/COLIPA study on UV filter chemicals. *ATLA* **27**, 247-259.

Cronin, M.T.D. & Dearden, J.C. (1995). QSAR in Toxicology 2. Prediction of acute mammalian toxicity and interspecies correlations. *Quantitative Structure-Activity Relationships* **14**, 117-120.

Curren, R., Bruner, L., Goldberg, A. & Walum, E. (1998). Validation and acute toxicity testing. Proceedings of the 13th meeting of the Scientific Group on Methodologies for the Safety Evaluation of Chemicals (SGOMSEC). *Environmental Health Perspectives* **106** (Suppl. 2), 419-425.

De Jongh, J., Forsby, A., Houston, J.B., Beckman, M., Combes, R. & Blauuboer, B.J. (1999). An integrated approach to the prediction of systemic toxicity using computer-based biokinetic models and biological *in vitro* test methods: overview of a prevalidation study based on the ECITTS project. *Toxicology in Vitro* 13, 549-554.

Delistraty, D., Taylor, B. & Anderson, R. (1998). Comparisons of acute toxicity of selected chemicals to rainbow trout and rats. *Ecotoxicology and Environmental Safety* **39**, 195-200.

ECVAM (1998a). Statement on the scientific validity of the 3T3 NRU PT test (an in vitro test for phototoxic potential). *ATLA* **26**, 7-8.

ECVAM (1998b). Statement on the application of the 3T3 NRU PT test to UV filter chemicals. *ATLA* **26**, 385-386.

Ekwall, B., Clemedson, C., Crafoord, B., Ekwall, Ba., Hallander, S., Walum, E. & Bondesson, I. (1998). MEIC Evaluation of Acute Systemic Toxicity. Part V. Rodent and Human Toxicity Data for the 50 Reference Chemicals. *ATLA* **26**, 571-616.

Ekwall, B., Clemedson, C., Ekwall, B., Ring, P. & Romert, L. (1999). EDIT: a new international multicentre programme to develop and evaluate batteries of *in vitro* tests for acute and chronic systemic toxicity. *ATLA* **27**, 339-349.

Ekwall, B., Ekwall, B. & Sjostrom, M. (2000). MEIC evaluation of acute systemic toxicity. Part VIII. Multivariate partial least squares evaluation, including the selection of a battery cell line tests with a good prediction of human acute lethal peak blood concentrations for 50 chemicals. *ATLA* **28**, 201-234.

Fentem, J., Fry, J., Garle, M., Gülden, M., Seibert, H., Voss, J-U., Wassermann, O., Perchermeier, M. & Wiebel, F. (1993). An international evaluation of selected *in vitro* toxicity test systems for predicting acute systemic toxicity. A report prepared for DGXI, CEC; Contract Numbers B92/B4-3063/14086 & B92/B4-3040/14087. FRAME Nottingham.

Fry, J.R., Garle, M.J., Hammond, A.H. & Hatfield, A. (1990). Correlation of acute lethal potency with *in vitro* cytotoxicity. *Toxicology in Vitro* **4**, 175-178.

Genschow, E., Scholz, G., Brown, N.A., Piersma, A.H., Brady, M., Clemann, N., Huuskonen, H., Paillard, F., Bremer, S. & Spielmann, H. (1999). Development of prediction models for three *in vitro* embryotoxicity tests which are evaluated in an ECVAM validation study. *ALTEX* **16**, 73-83.

Halle, W. & Goeres, E. (1988). Register der zytotoxizität (IC50) in der zellkultur und möglichkeiten zur abschätzung der akuten toxizität (LD50). In *Beiträge zur Wirkstofforschung, Institute für Wirkstofforschung* (ed. P. Oehme, H. Loewe & E. Goeres). pp 1-108. Berlin, Germany: Institut fuer Wirkstofforschung.

Halle, W., Spielmann, H. & Liebsch, M. (2000). Prediction of human lethal concentrations by cytotoxicity data from 50 MEIC chemicals. *ALTEX* **17**, 75-79.

Hulme, L.M., Reeves, H.L., Clothier, R.H., Smith, M. & Balls M. (1987). Assessment of two alternative methods for predicting the *in vivo* toxicities of metallic compounds. *Molecular Toxicology* **1**, 589-596.

Johnson, S.R. & Jurs, P.C. (1997). Prediction of acute mammalian toxicity from molecular structure for a diverse set of substituted anilines using regression analysis and computational neural networks. In *Computer-Assisted Lead Finding and Optimization* (ed. H. van de Waterbeemd, B. Testa & G. Folkers). pp 31-48. Basel: Verlag Helvetica Chimica Acta.

Knox, P., Uphill, P.F., Fry, J.R., Benford, J. & Balls, M. (1986). The FRAME multicentre project on *in vitro* cytotoxicology. *Food and Chemical Toxicology* **24**, 457-463.

OECD (1996). OECD Guideline for Testing of Chemicals. 423. Acute Oral Toxicity - Acute Toxic Class Method. Organisation for Economic Cooperation and Development, Paris. Updated in 2000: http://www.oecd.org/ehs/test/testlist.htm.

OECD (1998). OECD Guideline for Testing of Chemicals. 425. Acute Oral Toxicity - Up-and-Down Procedure. Paris: Organisation for Economic Cooperation and Development. Updated in 2000: http://www.oecd.org/ehs/test/testlist.htm.

OECD (1992). OECD Guideline for Testing of Chemicals. 420. Acute Oral Toxicity - Fixed Dose Method. Paris: Organisation for Economic Cooperation and Development. Updated in 2000: http://www.oecd.org/ehs/test/testlist.htm.

Ohno, T., Futamura, Y., Harihara, A. *et al.* (1998a). Validation study of five cytotoxicity assays by JSSAE I. Overview of the study and analyses of variations of ED50 values. *AATEX* **5**, 1-38.

Ohno, T., Futamura, Y., Harihara, A. *et al.* (1998b). Validation study of five cytotoxicity assays by JSSAE VI. Details of the LDH release assay. *AATEX* **5**, 99-118.

Ohno, T., Futamura, Y., Harihara, A. *et al.* (1998c). Validation study of five cytotoxicity assays by JSSAE VIII. Details of the neutral red uptake assay. *AATEX* **5**, 131-145.

Reader, S.J., Blackwell, V., O'Hara, R., Clothier, R., Griffen, G. & Balls, M. (1989). A vital dye release method for assaying the short-term cytotoxic effects of chemicals and formulations. *ATLA* **17**, 28-33.

Riddell, R.J., Panacer, D.S., Wilde, S.M., Clothier, R.H. & Balls, M. (1986a). The importance of exposure period and cell type in *in vitro* toxicity tests. *ATLA* **14**, 86-92.

Riddell, R.J., Clothier, R.H. & Balls, M. (1986b). Evaluation of three *in vitro* cytotoxicity assays. *Food and Chemical Toxicology* **24**, 469-471.

Seibert, H., Balls, M., Fentem, J.H., Bianchi, V., Clothier, R.H., Dierickx, P.J., Ekwall, B., Garle, M.J., Gómez-Lechón, M.J, Gribaldo, L., Gülden, M., Kiebsch, M., Rasmussen, E., Roguet, R., Shivrastava, R., & Walum, E. (1996). Acute toxicity testing *in vitro* and the classification and labelling of chemicals. The report and recommendations of ECVAM workshop 16. *ATLA* **24**, 499-510.

Scheers, E.M., Ekwall, B. & Dierickx, P.J. (2001). In vitro long-term cytotoxicity testing of 27 MEIC chemicals on Hep G2 cells and comparison with acute human toxicity data. *Toxicology in Vitro* **15**,153-161.

Spielmann, H., Balls, M., Dupuis, J., Pape W.J.W., Pechovitch, G., DeSilva, O., Hölzhutter, H.G., Clothier, R.H., Desolle, P., Gerberick, F., Liebsch, M., Lovell, W.W., Maurer, T., Pfannenbecker, U., Potthast, J.M., Csato, M., Sladowsly, D., Dteiling, W. & Brantom, P. (1998a). EU/COLIPA *In vitro* phototoxicity validation study, results of phase II (blind trial), part 1: the 3T3 NRU phototoxicity test. *Toxicology in Vitro* **12**, 305-327.

Spielmann, H., Balls, M., Dupuis, J., Pape W.J.W., DeSilva, O., Hölzhutter, H.G Gerberick, F., Liebsch, M., Lovell, W.W. & Pfannenbecker, U. (1998b). A study on the phototoxic potential of UV filter chemicals from Annex VII of EU Directive 76/768/EEC in the 3T3 NRU *in vitro* phototoxicity test. *ATLA* **26**, 679-705.

Spielmann, H., Genschow, E., Liebsch, M. & Halle, W. (1999). Determination of the starting dose for acute oral toxicity (LD50) testing in the up and down procedure (UDP) from cytotoxicity data. *ATLA* **27**, 957-966.

Start, D.M., Shopsis, C., Borenfreund, E. & Babich, H. (1986). Progress and problems in evaluating and validating alternative assays in toxicology. *Food and Chemical Toxicology* **24**, 449-455.

Triglia, D., Wegener., P.T. Harbell, J., Wallace, K., Matheson, D. & Shopsis, C. (1989). Interlaboratory validation study of the keratinocyte neutral red bioassay from Clonetics Corporation. In *Alternative Methods in Toxicology, Vol. 7* (ed. A.M. Goldberg). pp. 357-365. New York, USA: Mary Ann Liebert.

Wang, J., Lai, L. & Tang, Y. (1999). Data mining of toxic chemicals: Structure patterns and QSAR. *Journal of Molecular Modeling* **5**, 252-262.

Zbinden, G. & Flury-Roversi, M. (1981). Significance of the LD50 test for the toxicological evaluation of chemical substances. *Archives of Toxicology* **47**, 77-99.

Chapter 5. Local toxicity – acute dermal and ocular effects

Abraham, M.H., Kumarsingh, R., Cometto-Muniz, J.E. & Cain, W.S (1998a). A quantitative structure-activity relationship (QSAR) for a Draize eye irritation database. *Toxicology in Vitro* **12**, 201-207.

Abraham, M.H., Kumarsingh, R., Cometto-Muniz, J.E. & Cain, W.S (1998b). Draize eye scores and eye irritation thresholds in man combined into one quantitative structure-activity relationship. *Toxicology in Vitro* **12**, 403-408.

Barratt, M.D., Dixit, M.B., Jones, P.A. (1996). The use of *in vitro* cytotoxicity measurements in QSAR methods for the prediction of the skin corrosivity potential of acids. *Toxicology in Vitro* **10**, 283-290.

Barratt, M.D. (1995). A quantitative structure-activity relationship for the eye irritation potential of neutral organic chemicals. *Toxicology Letters* **80**, 69-74.

Barratt, M.D. (1996a). Quantitative structure-activity relationships for skin irritation and corrosivity of neutral and electrophilic organic chemicals. *Toxicology in Vitro* **10**, 247-256.

Barratt, M.D. (1996b). Quantitative structure-activity relationships (QSARs) for skin corrosivity of organic acids, bases and phenols: principal components and neural network analysis of extended datasets. *Toxicology in Vitro* **10**, 85-94.

Barratt, M.D., Brantom, P.G., Fentem, J.H., Gerner, I., Walker, A.P. & Worth, A.P. (1998). The ECVAM international validation study on *in vitro* tests for skin corrosivity. 1. Selection and distribution of the test chemicals. *Toxicology in Vitro* **12**, 471-482.

Basketter, D. A., Chamberlain, M., Griffiths, H. A., Rowson, M., Whittle, E. & York, M. (1997). The classification of skin irritants by human patch test. *Food and Chemical Toxicology* **35**, 845-852.

Borenfreund, E. & Puerner, J.A. (1985). Toxicity determined *in vitro* by morphological alterations and neutral red absorption. *Toxicology Letters* **24**, 119-124.

Botham, P. A., Earl, L. K., Fentem, J. H., Roguet, R. & van de Sandt, J. J. M. (1998) Alternative methods for skin irritation testing: the current status. ECVAM Skin Irritation Task Force Report 1. *ATLA* **26**, 195-211.

Bruner, L.H., de Silva, O., Earl, L.K., Easty, D.L., Pape, W. & Spielmann, H. (1998). Report on the COLIPA workshop on mechanisms of eye irritation. *ATLA* **26**, 811-820.

Burton, A.B.G., York, M. & Lawrence, R.S. (1981). The in vitro assessment of severe irritants. *Food and Cosmetics Toxicology* **19**, 471-480.

Cejkova, J., Lojda, Z., Brunova, B., Vacik, J. & Michalek, J. (1988). Disturbances in the rabbit cornea after short-term and long-term wear of hydrogel contact lenses. Usefulness of histochemical methods. *Histochemistry* **89**, 91-97.

Chetelat, A., Albertini, S., Dresp, J.H., Strobel, R. & Gocke, E. (1993a). Photomutagenesis test development: I. 8-Methoxypsoralen, chlorpromazine and sunscreen compounds in bacterial and yeast assays. *Mutation Research* **292**, 241-250.

Chetelat, A., Dresp, J.H. & Gocke, E. (1993b). Photomutagenesis test development: II. 8-Methoxypsoralen, chlorpromazine and sunscreen compounds in chromosomal aberration assays using CHO cells. *Mutation Research* **292**, 251-258.

Cronin, M.T.D. (1996). The use of cluster significance analysis to identify asymmetric QSAR data sets in toxicology. An example with eye irritation data. *SAR and QSAR in Environmental Research* **5**, 167-175.

Cronin, M.T.D., Basketter, D.A. & York, M. (1994). A quantitative structure-activity relationship (QSAR) investigation of a Draize eye irritation database. *Toxicology in Vitro* **8**, 21-28.

Dean, S.W., Lane, M., Dunmore, R.H., Ruddock, S.P., Martin, C.N., Kirkland, D.J. & Loprieno N. (1991). Development of assays for the detection of photomutagenicity of chemicals during exposure to UV light-I. Assay development. *Mutagenesis* **6**, 335-341.

EC (2000). Annex I to Commission Directive 2000/33/EC adapting to technical progress for the 27th time Council Directive 67/548/EEC on the approximation of laws, regulations and administrative provisions relating to the classification, packaging and labeling of dangerous substances. *Official Journal of the European Communities* **L136**, 91-97.

ECVAM (1998). ECVAM News & Views. *ATLA* **26**, 275-280.

ECVAM (2000). Statement on the application of the Epiderm™ human skin model for skin corrosivity testing. *ATLA* **28**, 365-366.

ECVAM (2001). Statement on the application of the CORROSITEX® assay for skin corrosivity testing. *ATLA* **29**, 96-97.

Fentem, J. H., Archer, G. E. B., Balls, M., Botham, P. A., Curren, R. D., Earl, L. K., Esdaile, D. J., Holzhütter, H. G. & Liebsch, M. (1998). The ECVAM international validation study on *in vitro* tests for skin corrosivity. 2. Results and evaluation by the Management Team. *Toxicology in Vitro* **12**, 483-524.

Fentem, J. H., Briggs, D., Chesné, C., Elliott, G. R., Harbell, J. W., Heylings, J. R., Portes, P., Roguet, R., van de Sandt, J. J. M. & Botham, P. A. (2001). A prevalidation study on *in vitro* tests for acute skin irritation: results and evaluation by the Management Team. *Toxicology in Vitro* **15**, 57-93.

Gautheron, P., Dukic, M., Alix, D. & Sina, J.F. (1992). Bovine corneal opacity and permeability test: an *in vitro* assay of ocular irritancy. *Fundamental and Applied Toxicology* **8**, 442-449.

Gerner, I., Graetschel, G., Kahl, J & Schlede, E. (2000a). Development of a decision support system for the introduction of alternative methods into local irritancy/corrosivity testing strategies. Development of a relational database. *ATLA* **28**, 11-28.

Gerner, I., Zinke, S., Graetschel, G. & Schlede, E. (2000b). Development of a decision support system for the introduction of alternative methods into local irritancy/corrosivity testing strategies. Creation of fundamental rules for a decision support system. *ATLA* **28**, 665-98.

Gettings, S.D., Bagley, D.M., Demetrulias, J.L., Dipasquale, L.C., Hintze, K.L., Rozen, M.G., Teal, J.J., Weise, S.L., Chudkowski, M., Marenus, K.D., Pape, W.J.W., Roddy, M.T., Schnetzinger, R., Silber, P.M., Glaza, S.M. & Kurtz, P.J. (1991). The CTFA evaluation alternatives program: an evaluation of *in vitro* alternatives to the Draize primary eye irritation test (Phase I). Hydro-alcoholic formulations; (Part 2) Data analysis and biological significance. *In Vitro Toxicology* **4**, 247-288.

Gettings, S.D., Bagley, D.M., Chudlowski, M., Demetrulias, J.L., Dipasquale, L.C., Galli, C.L., Gay, R., Hintze, K.L., Janus, J., Marenus, K.D., Muscatiello, M.J., Pape, W.J.W., Renskers, K.J., Roddy, M.T. & Schnetzinger, R. (1992). Development of potential alternatives to the Draize eye test. The CTFA evaluation of alternatives program (Phase II). Review of materials and methods. *ATLA* **20**, 164-171.

Gettings, S.D., Bagley, D.M., Casterton, M., Chudkowski, M., Curren, R.D., Demetrulias, J.L., Dipasquale, L.C., Feder, P.I., Galli, C.L., Gay, R., Glaza, S.M., Hintze, K.L., Janus, J., Kurtz, P.J., Lordo, R.A., Marenus, K.D., Moral, J., Muscatiello, M.J., Pape, W.J.W., Renskers, K.J., Roddy, M.T & Rozen, M.G. (1994). The CTFA evaluation of alternatives program: an evaluation of *in vitro* alternatives to the Draize primary eye irritation test (Phase II). Oil/water emulsions. *Food and Chemical Toxicology* **32**, 943-976.

Gettings, S.D., Lordo, R.A., Hintze, K.L., Bagley, D.M., Casterton, P.L., Chudkowski, M., Curren, R.D., Demetrulias, J.L., Dipasquale, L.C., Earl, L.K., Feder, P.I., Galli, C.L., Glaza, S.M., Gordon, V.C., Janus, J., Kurtz, P.J., Marenus, K.D., Moral, J., Pape, W.J., Renskers, K.J., Rheins, L.A., Roddy, M.T., Rozen, M.G., Tedeschi, J.P. & Zyracki, J. (1996). The CTFA Evaluation of Alternatives Program: an evaluation of *in vitro* alternatives to the Draize primary eye irritation test. (Phase III). Surfactant-based formulations. *Food and Chemical Toxicology* **34**, 79-117.

Hagino, S., Itagaki, H., Kato, S., Kobayashi, T. & Tanaka, M. (1991). Quantitative evaluation to predict the eye irritancy of chemicals: modification of chorioallantoic membrane test by using Trypan Blue. *Toxicology in Vitro* **5**, 301-304.

Hayashi, H., Nakamura, Y., Higashi, K., Kato, H., Kishida, F. & Kaneko, H. (1999). A quantitative structure-Activity relationship study of the skin irritation potential of phenols. *Toxicology in Vitro* **13**, 915-922.

Heylings, J. R. Clowes, H.M. & Hughes L. (2001). Comparison of tissue sources for the skin integrity function test (SIFT). *Toxicology in Vitro* **13**, 597-600.

Kulkarni, A.S. & Hopfinger, A.J. (1999). Membrane-interaction QSAR analysis: application to the estimation of eye irritation by organic compounds. *Pharmaceutical Research* **16**, 1245-1253.

Liebsch, M., Traue, D., Barrabas, C., Spielmann, H., Uphill, P., Wilkins, S., Wiemann, C., Kaufmann, T., Remmele, M. & Holzhütter, H. G. (2000). The ECVAM prevalidation study on the use of EpiDerm for skin corrosivity testing. *ATLA* **28**, 371-401.

Lüpke, N.P. (1985). Hen's egg chorioallantoic membrane test for irritation potential. *Food & Chemical Toxicology* **23**, 287-291.

Mosmann, T. (1983). Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays. *Journal of Immunoogical Methods* **65**, 55-63.

NIH (1999). *Corrositex®: an in vitro test method for assessing dermal corrosivity potential of chemicals*. NIH Publication No. 99-4495. Research Triangle Park, NC, USA: NIEHS.

Nussenblatt, R.B., Bron, A., Chambers, W., McCulley, J.P., Pericoi, M. & Ubels, J.L. (1998). Ophthalmologic Perspectives on Eye Irritation Testing. *Journal of Toxicology - Cutaneous & Ocular Toxicology* **17**,103-109.

OECD (2001). *Revised proposals for updated test guidelines 404 and 405: dermal and eye corrosion/irritation studies.* ENV/JM/TG(2001)2, 112 pp. Paris, France: OECD.

Patlewicz, G.Y., Rodford, R.A., Ellis, G. & Barratt, M.D. (2000). A QSAR model for the eye irritation of cationic surfactants. *Toxicology in Vitro* **14**, 79-84.

Pape, W.J.W. & Hoppe, U. (1991). *In vitro* methods for the assessment of primary local effects of topically applied preparations. *Skin Pharmacology* **4**, 205-212.

Reader, S.J., Blackwell, V., O'Hara, R., Clothier, R., Griffin, G. & Balls, M. (1989). A vital dye release method for assaying the short-term cytotoxic effects of chemicals and formulations. *ATLA* **17**, 28-33.

Robinson, M. K., Osborne, R. & Perkins, M. A. (2000). *In vitro* and human testing strategies for skin irritation. *Annals of the New York Academy of Sciences* **919**, 192-204.

Rosenkranz, H.S., Zhang, Y.P. & Klopman, G. (1998). The development and characterisation of a structure-activity relationship model of the Draize eye irritation test. *ATLA* **26**, 779-809.

Stern, M., Klausner, M., Alvarado, R., Renskers, K. & Dickens, M. (1998). Evaluation of the EpiOcular^{TM} tissue model as an alternative to the Draize eye irritation test. *Toxicology in Vitro* **12**, 455-461.

Smith, J.S., Macina, O.T., Sussman, N.B., Luster, M.I. & Karol, M.H. (2000). A robust structure-activity relationship (SAR) model for esters that cause skin irritation in humans. *Toxicological Sciences* **55**, 215-222.

Spielmann, H., Gerner, I., Kalweit, S., Moog, R., Wirnserberger, T., Krauser, K., Kreiling, R., Kreuzer, H., Luepke, N.P., Miltenburger, H.G., Müller, N., Mürmann, P., Pape, W., Siegmund, B., Spengler, J., Steiling, W. & Wiebel, F.J. (1991). Interlaboratory assessment of alternatives to the Draize eye irritation test in Germany. *Toxicology in Vitro* **5**, 539-542.

Spielmann, H., Kalweit, S., Liebsch, M., Wirnserberger, T., Gerner, I., Bertram-Neis, E., Krauser, K., Kreiling, R., Miltenburger, H.G., Pape, W. & Steiling, W. (1993). Validation study of alternatives to the Draize eye irritation test in Germany: cytotoxicity testing and HET-CAM test with 136 industrial chemicals *Toxicology in Vitro* **7**, 505-510.

Spielmann, H., Lovell, W.W., Hölzle, E., Johnson, B.E., Maurer, T., Miranda, M.A., Pape, W.J.W, Sapora, O. & Sladowski, D. (1994). *In vitro* phototoxity testing. The report and recommendations of ECVAM workshop 2. *ATLA* **22**, 314-348.

Spielmann, H., Kalweit, S., Liebsch, M., Moldenhauer, F., Wirnserberger, T., Hölzhutter, H.G., Schneider, B., Glaser, S., Gerner, I., Pape, W., Kreiling, R., Krauser, K., Miltenburger, H.G., Steiling, W., Luepke, N.P., Müller, N., Kreuzer, H., Mürmann, P., Spengler, J., Bertram-Neis, E., Siegmund, B. & Wiebel, F.J. (1996). Results of a validation study in Germany on two *in vitro* alternatives to the Draize eye irritation test, the HET-CAM test and the 3T3 NRU cytotoxicity test. *ATLA* **24**, 741-858.

Spielmann, H., Balls, M., Dupuis, J., Pape W.J.W., Pechovitch, G., DeSilva, O., Hölzhutter, H.G., Clothier, R.H., Desolle, P., Gerberick, F., Liebsch, M., Lovell,

W.W., Maurer, T., Pfannenbecker, U., Potthast, J.M., Csato, M., Sladowsly, D., Dteiling, W. & Brantom, P. (1998a). EU/COLIPA *In vitro* phototoxicity validation study, results of phase II (blind trial), part 1: the 3T3 NRU phototoxicity test. *Toxicology in Vitro* **12**, 305-327.

Spielmann, H., Balls, M., Dupuis, J., Pape W.J.W., DeSilva, O., Hölzhutter, H.G Gerberick, F., Liebsch, M., Lovell, W.W. & Pfannenbecker, U. (1998b). A study on the phototoxic potential of UV filter chemicals from Annex VII of EU Directive 76/768/EEC in the 3T3 NRU *in vitro* phototoxicity test. *ATLA* **26**, 679-705.

Spielmann, H., Müller, L., Averbeck, D., Balls, M., Brendler-Schwaab, S., Castell, J.V., Curren, R., de Silva, O., Gibbs, N.K., Liebsch, M., Lovell, W.W., Merk, H.F., Nash, J.F., Neumann, N.J., Pape, W.J.W., Ulrich, P. & Vohr, H-W. (2000). The second ECVAM workshop on phototoxicity testing. The report and recommendations of ECVAM workshop 42. *ATLA* **28**, 777-814.

Tchao, R. (1988). Trans-epithelial permeability of fluorescein *in vitro* as an assay to determine eye irritants. *In Alternative Methods in Toxicology, Volume 6, Progress in In Vitro Toxicology* (ed. A.M. Goldberg). pp. 271-283. New York: Mary Ann Liebert.

Wallin, R.F., Hume, R.D. & Jackson, E.M. (1987). The agarose diffusion method for ocular irritancy screening cosmetic products, part I. Journal of Toxicology. *Cutaneous and Ocular Toxicology* 6, 239-250.

Whittle, E.G., Barratt, M.D., Carter, J.A., Basketter, D.A. & Chamberlain, M. (1996). The skin corrosivity potential of fatty acids: *in vitro* rat and human skin testing and QSAR studies. *Toxicology in Vitro* **10**, 95-100.

Whittle, E., Basketter, D., York, M., Kelly, L., Hall, , T., McCall, J., Botham, P., Esdaile, D. & Gardner, J. (1992). Findings of an interlaboratory trial of the enucleated eye method as an alternative eye irritation test. *Journal of Pharmacological and Toxicological Methods* **2**, 30-41.

Worth, A.P. (2000). The integrated use of physicochemical and *in vitro* data for predicting chemical toxicity. PhD Thesis. Liverpool John Moores University, UK.

Worth, A.P. & Cronin, M.T.D. (1999). Embedded cluster modelling – a novel method for analysing embedded data sets. *Quantitative Structure-Activity Relationships* **18**, 229-235.

Worth A.P. & Cronin, M.T.D. (2000). Embedded cluster modelling: a novel QSAR method for generating elliptic models of biological activity. In *Progress in the Reduction, Refinement and Replacement of Animal Experimentation* (eds. M. Balls, A-M. van Zeller & M.E. Halder). pp. 479-491. Elsevier Science, Amsterdam.

Worth, A.P. & Cronin, M.T.D. (2001a). The use of bootstrap resampling to assess the uncertainty of Cooper statistics. *ATLA* **29**, 447-459.

Worth, A.P. & Cronin, M.T.D. (2001b). The use of pH measurements to predict the potential of chemicals to cause acute dermal and ocular toxicity. *Toxicology* **169**, 119-131.

Worth, A.P. & Cronin, M.T.D. (2001c). Prediction models for eye irritation potential based on endpoints of the HETCAM and neutral red assays. *In Vitro & Molecular Toxicology* **14**, 143-156.

Worth A.P. & Fentem J.H. (1999). A general approach for evaluating stepwise testing strategies. *ATLA* **27**, 161-177.

Zinke, S., Gerner, I., Graetschel, G. & Schlede, E. (2000). Local irritation/corrosion testing strategies: development of a decision support system for the introduction of alternative methods. *ATLA* **28**, 29-40.

Zuang, V. Balls, M., Botham, P.A., Coquette, A., Corsini, E., Curren, R.D., Elliott, G.R., Fentem, J.H., Heylings, J.R., Liebsch, M., Medina, J., Roguet, R., van de Sandt, H., Wiemann, C. & Worth, A.P. (2002). Follow-up to the ECVAM Prevalidation Study on *In Vitro* Tests for Acute Skin Irritation. ECVAM Skin Irritation Task Force Report 2. *ATLA* **30**, 109-129.

Chapter 6. Local toxicity - sensitisation

Barratt, M.D., Basketter, D.A., Chamberlain, M., Admans, G.D. & Langowski, J.J. (1994). An expert system rulebase for identifying contact allergens. *Toxicology in Vitro* **8**, 1053–1060.

Barratt, M.D. (1995). Quantitative structure-activity relationships for skin permeability. *Toxicology in Vitro* **9**, 27-37.

Barratt, M.D. & Langowski, J.J. (1999). Validation and subsequent development of the DEREK skin sensitization rulebase by analysis of the BgVV list of contact allergens. *Journal of Chemical Information and Computer Sciences* **9**, 294–298.

Cronin, M.T.D & Basketter, D.A. (1994). Multivariate QSAR analysis of a skin sensitization database. *SAR and QSAR in Environmental Research* **2**, 159-179.

Dearman, R.J., Smith, S., Basketter, D.A. & Kimber, I. (1996). Classification of chemical allergens according to cytokine secretion profiles of murine lymph node cells. *Journal of Applied Toxicology* **17**, 53-62.

Dearman, R.J. & Kimber, I. (2001). Cytokine fingerprinting and hazard assessment of chemical respiratory allergy. *Journal of Applied Toxicology* **21**, 153-163.

ECVAM (2000). Statement on on the validity of the local lymph node assay for skin sensitisation testing. *ATLA* **28**, 366-367.

Enslein, K., Gombar, V.K., Blake, B.W., Maibach, H.I., Hostynek, J.J., Sigman, C.C. & Bagheri, D. (1997). A quantitative structure-toxicity relationships model for the

dermal sensitization guinea pig maximization assay. *Food & Chemical Toxicology* **35**, 1091-1098.

Gerberick, G.F., Ryan, C.A., Kimber, I., Dearman, R.J., Lea, L.J. & Basketter, D.A. (2000). Local lymph node assay: validation assessment for regulatory purposes. *The American Journal Contact Dermatology* **11**, 3-18.

Graham, C., Gealy, R., Macina, O.T., Karol, M.H. & Rosenkranz, H.S. (1996). QSAR for allergic contact dermatitis. *Quantative Structure-Activity Relationships* **15**, 224-229.

Kimber, I. (2000). Skin sensitisation: immunological mechanisms and novel approaches to predictive testing. In *Progress in the Reduction, Refinement, and Replacement of Animal Experimentation* (ed. M. Balls, A.-M. van Zeller and M.E. Halder). pp. 613-622. Amsterdam: Elsevier Science.

Kimber, I., Pichowski, J.S., Betts, C.J., Cumberbatch, M., Basketter, D.A. & Dearman, R.J. (2001). Alternative approaches to the identification and characterisation of chemical allergens. *Toxicology in Vitro* **15**, 307-312.

NIH (1999). The murine local lymph node assay. The results of an independent peer review evaluation coordinated by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) and the National Toxicology Program Center for the Evaluation of Alternative Toxicological Methods (NICEATM). NIH Publication no. 99-4494. 211 pp. Research Triangle Park, NC, USA: NIH.

OECD (2001). *Revised proposal for new test guidelines 429: skin sensitisation: local lymph node assay.* ENV/JM/TG(2001)4, 65 pp. Paris, France: OECD.

Pichowski, J., Cumberbatch, M., Dearman, R.J., Basketter D.A. & Kimber, I. (2000). Investigation of induced changes in interleukin 1beta mRNA expression by cultured human dendritic cells as an in vitro approach to skin sensitization testing. *Toxicology in Vitro* **14**, 351-360.

Régnier, M., Patwardhan, A., Scheynius, A. & Schmidt, R. (1998). Reconstructed human epidermis composed of keratinocytes, melanocytes and Langerhans cells. *Medical & Biological Engineering & Computing* **36**, 821-824.

Ridings, J.E., Barratt, M.D., Cary, R., Earnshaw, C.G., Eggington, C.E., Ellis, M.K., Judson, P.N., Langowski, J.J., Marchant, C.A., Payne, M.P., Watson, W.P. & Yih, T.D. (1996). Computer prediction of possible toxic action from chemical structure: an update on the DEREK system. *Toxicology* **106**, 267–279.

Ramu, K., Perry, C.S., Ahmed, T., Pakenham, G. & Kehrer, J.P. (1996). Studies on the basis for the toxicity of acrolein mercapturates. *Toxicology & Applied Pharmacology* **140**, 487-498.

Ryan, C.A., Hulette, B.C. & Gerberick, G. (2001). Approaches for the development of cell-based in vitro methods for contact sensitization. *Toxicology in Vitro* **15**, 43-55.

Sanderson, D.M. & Earnshaw, C.G. (1991). Computer prediction of possible toxic action from chemical structure: The DEREK system. *Human and Experimental Toxicology* **10**, 261–273.

Tuschl, H., Kovac, R. & Weber, E. (2000). The expression of surface markers on dendritic cells as indicators for the sensitizing potential of chemicals. *Toxicology in Vitro* **14**, 541-549.

Warbrick, E.V., Dearman, R.J. & Kimber, I. (2002). Induced changes in total serum IgE concentration in the Brown Norway rat: potential for identification of chemical respiratory allergens. *Journal of Applied Toxicology* **22**, 1-11.

Zhu, J., Rogers, A.V., Burke-Gaffney, A., Hellewell, P.G. & Jeffery, P.K. (1999). Cytokine-induced airway epithelial ICAM-1 upregulation: quantification by high-resolution scanning and transmission electron microscopy. *European Respiratory Journal* **13**, 1318-1328.

Chapter 7. Biokinetics

Computer modelling

Abraham, M.H., Martins, F. & Mitchell, R.C. (1997). Algorithms for skin permeability using hydrogen bond descriptors: the problem of steroids. *Journal of Pharmacy and Pharmacology* **49**, 858-865.

Abraham, M.H., Chadha, H.S., Martins, F., Mitchell, R.C., Bradbury, M.W. & Gratton, J.A. (1999). Hydrogen bonding part 46: a review of the correlation and prediction of transport properties by an LFER method: physico-chemical properties, brain penetration and skin permeability. *Pesticide Science* **55**, 78-88.

Agatonovic-Kustrin, S., Tucker, I.G., Zecevic, M. & Zivanovic, L.J. (2000). Prediction of drug transfer into human milk from theoretically derived descriptors. *Analytica Chimica Acta* **418**, 181-195.

Andersen, M.E. (1991). Physiological modelling of organic compounds. *Annals of Occupational Hygiene* **35**, 309-321.

Archer, G., Balls, M., Bruner, L.H., Curren, R.D., Fentem, J.H., Holzhütter, H-G., Liebsch, M., Lovell, D.P. & Southee, J.A. (1997). The validation of toxicological prediction models. *ATLA* **25**, 505-516.

Barratt, M.D. (1995). Quantitative structure-activity relationships for skin permeability. *Toxicology in Vitro* **9**, 27-37.

Blaauboer, B.J., Forsby, A., Houston, J.B., Beckman, M., Combes, R.D. & DeJongh, J. (2000). An integrated approach to the prediction of systemic toxicity using biokinetic models and biological *in vitro* test methods. In *Progress in the Reduction, Refinement and Replacement of Animal Experimentation* (eds. M. Balls, A-M. van Zeller & M.E. Halder). pp. 525-536. Elsevier Science, Amsterdam.

Brown, R.P., Foran J., Olin S. & Robinson, D. (1994*). Physiological Parameter Values for PBPK Models.* 100 pp. Washington, DC: International Life Sciences Institute (ILSI) Risk Science Institute (RSI).

Clark, D.E. & Pickett, S.D. (2000). Computational methods for the prediction of 'drug-likeness'. *Drug Discovery Today* **5**, 49-58.

Clewell, H.J. & Andersen, M.E. (1986). A multiple dose-route physiological model for volatile chemicals using ACSL/PC. In *Languages for Continuous simulation* (ed. F.D. Cellier). San Diego, USA: Society for computer simulation.

Cronin, M.T.D, Dearden, J.C., Worth, A.P. & Walker, J.D. (2002). QSARs for predicting environmental-human health interactions. *Environmental Toxicology and Chemistry*, in press.

Cronin, M.T.D., Dearden, J.C. & Moss, G.P. (2002). QSARs for the penetration of chemicals across human, fish and artificial membranes – mechanistic considerations. In *Quantitative Structure-Activity Relationships in Environmental Sciences -VIII* (ed. J.D. Walker). Pensacola, FL, USA: SETAC Press, in press.

Cronin, M.T.D., Dearden, J.C., Gupta, R. & Moss, G.P. (1998). An investigation of the mechanism of flux across polydimethylsiloxane membranes by quantitative structure-permeability relationships. *Journal of Pharmacy and Pharmacology* **50**, 143-152.

Cronin, M.T.D, Dearden, J.C., Moss, G.P. & Murray-Dickson, G. (1999). Investigation of the mechanism of flux across human skin *in vitro* by quantitative structure-permeability relationships. *European Journal of Pharmaceutical Sciences* **7**, 325-330.

DeJongh, J., Verhaar, H.J. & Hermens, J.L. (1997). A quantitative property-property relationship (QPPR) approach to estimate *in vitro* tissue-blood partition coefficients of organic chemicals in rats and humans. *Archives of Toxicology* **72**, 17-25.

DeJongh, J., Verhaar, H.J.M., Hermens, J.L.M. (1998) Role of kinetics in acute lethality of nonreactive volatile organic compounds (VOCs). *Toxicological Science* **45**, 26-32.

Duffy, J.C., Cronin, M.T.D., Dearden, J.C., Moss, G.P. & Murray-Dickson, G. (1999). Prediction of blood-brain barrier partitioning using QSAR analysis. *Journal of Pharmacy and Pharmacology* **51** (Supplement), 260.

Egan, W.J., Merz, K.M. & Baldwin, J.J. (2000). Prediction of drug absorption using multivariate statistics. *Journal of Medicinal Chemistry* **43**, 3867-3877.

Feher, M., Sourial, E. & Schmidt, J.M. (2000). A simple model for the prediction of blood-brain partitioning. *International Journal of Pharmaceutics* **201**, 239-247.

Fiserova-Bergerova, V. & Diaz, M.L. (1986). Determination and prediction of tissue-gas partition coefficients. *International Archives of Occupational and Environmental Health* **58**, 75-87.

Franke, H., Galla, H. & Beuckmann, C.T. (2000). Primary cultures of brain microvessel endothelial cells: a valid and flexible model to study drug transport through the blood-brain barrier *in vitro*. *Brain Research Protocols* **5**, 248-256.

Houston, J.B. & Carlile, D.J. (1987). Prediction of hepatic clearance from microsomes, hepatocytes and liver slices. *Drug Metabolism Reviews* **29**, 891-922.

Ito, K., Iwatsubo, T., Kanamitsu, S., Nakajima, Y. & Sugiyama, Y. (1998). Quantitative prediction of *in vivo* drug clearance and drug interactions from *in vitro* data on metabolism, together with binding and transport. *Annual Review of Pharmacology and Toxicology* **38**, 461-499.

Jepson, G.W., Hoover, D.K, Black, R.K., McCafferty, J.D., Mahle, D.A. & Gearhart, J.M. (1994). A partition coefficient determination method for nonvolatile chemicals in biological tissues. *Fundamental and Applied Toxicology* **22**, 519-524.

Kirchner, L.A., Moody, R.P., Doyle, E., Bose, R., Jeffery, J. & Chu, I. (1997). The prediction of skin permeability by using physicochemical data. *ATLA* **25**, 359-370.

Lipinski, C.A., Lombardo, F., Dominy, B.W. & Feeney, P.J. (1997). Experimental and computational approaches to estimate solubility and permeability in drug discovery and devlopment settings. *Advanced Drug Delivery Reviews* **23**, 3-25.

Moss, G.P., Dearden, J.C., Patel, H. & Cronin, M.T.D. (2002). Quantitative Structure Permeability Relationships (QSPRs) for Percutaneous Absorption. *Toxicology in Vitro*, in press.

Österberg, T. & Norinder, U. (2000). Prediction of polar surface area and drug transport processes using simple parameters and PLS statistics. *Journal of Chemical Information and Computer Sciences* **40**, 1408-1411.

Pacifini, G.M. & Nottoli, R. (1995). Placental transfer of drugs administered to the mother. *Clinical Pharmacokinetics* **28**, 235-269.

Pacifini, G.M. & Viani, A. (1992). Methods of determining plasma and tissue binding of drugs. *Clinical Pharmacokinetics* **23**, 449-468.

Perbellini, L., Brugnone, F., Caretta, D. & Maranelli, G. (1985). Partition Coefficients of some industrial aliphatic hydrocarbons (C5-C7) in blood and human tissues. *British Journal of Industrial Medicine* **42**, 162-167.

Platts, J.A., Abraham, M.H., Hersey, A. & Butina, D. (2000). Estimation of molecular linear free energy relationship descriptors. 4. Correlation and prediction of cell permeation. *Pharmaceutical Research* **17**, 1013-1018.

Sato, A. & Nakajima, T. (1979). A structure-activity relationship of some chlorinated hydrocarbons. *Archives of Environmental Health* **34**, 69-75.

Stenberg, P., Norinder, U., Luthman, K. & Artursson, P. (2001). Experimental and computational screening models for the prediction of intestinal drug absorption. *Journal of Medicinal Chemistry* **44**, 1927-1937.

Tice, T.M. (2001). Selecting the right compounds for screening: does Lipinski's Rule of 5 for pharmaceuticals apply to agrochemicals? *Pesticide Management Science* **57**, 3-16.

van de Waterbeemd, H., Smith, D.A. & Jones, B.C. (2001). Lipophilicity in PK design: methyl, ethyl, futile. *Journal of Computer-Aided Molecular Design* **15**, 273-286.

Worth, A.P. & Balls, M. (2001). The importance of the prediction model in the development and validation of alternative tests. *ATLA* **29**, 135-143.

Worth, A.P., Barratt, M.D. & Houston, J.B. (1998). The validation of computational prediction techniques. *ATLA* **26**, 241-247.

Worth, A.P. & Cronin M.T.D. (2000). Structure-permeability relationships for transcorneal penetration. *ATLA* **28**, 403-413.

Zhao, Y.H., Le, J., Abraham, M.H., Hersey, A., Eddershaw, P.J., Luscombe, C.N., Boutina, D., Beck, G., Sherborne, B., Cooper, I. & Platts, J.A. (2001). Evaluation of human intestinal absorption data and subsequent derivation of a quantitative structure-activity relationship (QSAR) with the Abraham descriptors. *Journal of Pharmaceutical Sciences* **90**, 749-784.

Percutaneous absorption

Bronaugh, R.L. & Stewart, R.F. (1984). Methods for *in vitro* percutaneous absorption studies. III. Hydrophobic compounds. *Journal of Pharmaceutical Sciences* **73**, 1255-1259.

Diembeck, W, Beck, H, Benech-Kieffer, F, Courtellemont, P, Dupuis, J, Lovell, W, Paye, M, Spengler, J, Steiling, W. (1999). Test Guidelines for In Vitro Assessment of Dermal Absorption and Percutaneous Penetration of Cosmetic Ingredients. *Food and Chemical Toxicology* **37**, 191-205.

ECETOC (1993). *Percutaneous Absorption*. ECETOC Monograph No. 20. Brussels: European Centre for Ecotoxicology and Toxicology of Chemicals.

Howes, D., Guy, R., Hadgraft, J., Heylings, J., Hoeck, U., Kemper, F., Maibach, H., Marty, J-P., Merk, H., Parra, J., Rekkas, D., Rondelli, I., Schaefer, H., Täuber, U. & Verbiese, N. (1996). Methods for assessing percutaneous absorption. The report and recommendations of ECVAM Workshop 13. *ATLA* **24**, 81-106.

OECD (2000a). Draft Test Guideline 428. Skin absorption: in vitro Method. Paris: Organisation for Economic Cooperation and Development.

OECD (2000b). Draft Test Guideline 427. Skin absorption: in vivo Method. Paris: Organisation for Economic Cooperation and Development.

OECD (2000c). Draft Guidance Document for the Conduct of Skin Absorption Studies. Paris: Organisation for Economic Cooperation and Development.

Scott, R.C., Batten, P.L., Clowes, H.M., Jones, B.K. & Ramsey, J.D. (1992). Further validation of an *in vitro* method to reduce the need for *in vivo* studies for measuring the absorption of chemicals through the skin. *Fundamental and Applied Toxicology* **19**, 484-492.

GI barrier

Ayrton, A. & Morgan, P. (2001). Role of transport proteins in drug absorption, distribution and excretion. *Xenobiotica* **31**, 469-407.

Hidalgo, I.J., Raub, T.J. & Borchardt, R.T. (1989). Characterization of the human colon carcinoma cell line (Caco-2) as a model system for intestinal epithelial permeability. *Gastroenterology* **96**, 736-49.

Le Ferrec, E., Chesne, C., Artusson, P., Brayden, D., Fabre, G., Gires, P., Guillou, F., Rousset, M., Rubas, W. & Scarino, M.L. (2001). *In vitro* models of the intestinal barrier. The report and recommendations of ECVAM Workshop 46. *ATLA* **29**, 649-668.

Lennernas, H. (1998). Human intestinal permeability. *Journal of Pharmaceutical Sciences* **87**, 403-410.

Li, A.P. (2001). Screening for human ADME/Tox drug proteins in drug discovery. *Drug Discovery Today* 6, 357-366.

Li, A.P.S. et al. (DATE ?). Caco-2 cell screening assay for intestinal absorption. In *High throughput screening: Supplement to Biochemical Products (January)*. pp. 6-9. Cahners.

Stanley, S.D. & Wilding, I.R. (2001). Oral drug absorption studies: the test model for man is man!. *Drug Discovery Today* **6**, 127-128.

Taipalensuu, J. et al. (2001). Correlation of gene expression of ten drug efflux proteins of the ATP-binding cassette transporter family in normal human jejunum and in human intestinal epithelial caco-2 cell monolayers. *Journal of Pharmacology and Experimental Therapeutics* **299**, 164-170.

Ungell, A-L. (1997). In vitro absorption studies and their relevance to absorption from the GI tract. *Drug Development and Industrial Pharmacy* **23**, 879-892.

BBB

Cestelli, A., Catania, C., D'Agostino, S., Di Liegro, I., Licata, L., Schiera, G., Pitarresi, G.L., Savettieri, G., De Caro, V., Giandalia, G., Giannola, L.I. (2001). Functional feature of a novel model of blood-brain barrier: studies on permeation of test compounds. *Journal of Controlled Release* **76**, 139-147.

Foster, K.A. & Roberts, M.S. (2000). Experimental methods for studying drug uptake in the head and brain. *Current Drug Metabolism* **1**, 333-356.

Garberg, P. (1998). *In vitro* models of the blood-brain barrier. *ATLA* **26**, 821-847.

Kusuhara, H. & Sugiyama, Y. (2001). Efflux transport systems for drugs at the blood-brain barrier and blood-fluid-cerebrospinal fluid barrier (part 1). *Drug Discovery Today* **6**, 150-156.

Stanness, K.A., Westrum, L.E., Fornaciari, E., Mascagni, P., Nelson, J.A., Stenglein, S.G., Myers, T. & Janigro, D. (1997). Morphological and functional characterization of an in vitro blood-brain barrier model. *Brain Research* **771**, 329-342.

Veronesi, B. & Ehrich, M. (1999). Cell culture models of environmental chemical neurotoxicity. In *Neurotoxicology In Vitro* (ed. V.W. Pentreath). pp. 239-268. London: Taylor and Francis.

Metabolism

Brown, S.J., Raja, A.A. & Lewis, D.F.V. (1994). A comparison between COMPACT and Hazard Expert evaluations for 80 chemicals tested by the NTP/NCI rodent bioassay. *ATLA* **22**, 482-500.

Bull, S., Langezaal, I., Clothier, R. & Coecke, S. (2001). A genetically engineered cell-based model system for detecting metabolism-mediated toxicity. *ATLA* **29**, 703-716.

Burchell, B., Brierley, C.H., Monaghan, G. & Clarke, D.J. (1998). The structure and function of the UDP-glucuronosyltransferase gene family. *Advances in Pharmacology* **42**, 335-338.

Burchell, B., Soars, M., Monaghan, G., Cassidy, A., Smith, D. & Ethell, B. (2000). Drug-mediated toxicity caused by genetic deficiency of UDP-glucuronosyltransferases. *Toxicology Letters* **112-113**, 333-340.

Clarke, S.E. (1998). In vitro assessment of human cytochrome P450. *Xenobiotica* **28**, 1167-1202.

Coecke, S., Catalani, P., Bull, S., Langezaal, I., Bogni, A., Bremer, S. & Balls, M. (2000). The integrated use of a genetically engineered cell line panel harbouring drug metabolising enzymes to detect metabolism-related problem compounds. In *Progress in the Reduction, Refinement, and Replacement of Animal Experimentation* (ed. M. Balls, A.-M. van Zeller and M.E. Halder). pp. 147-157. Amsterdam: Elsevier Science.

Coecke, S., Bogni, A., Worth, A., Hartung, T. & Monshouwer, M. (2001). The use of genetically engineered cells for assessing CYP2D6-related polymorphic effects. *Toxicology in Vitro* **15** 553-556.

Coecke, S., Rogiers, V., Bayliss, M, Castell, J., Doehmer, J., Fabre, G., Fry, J., Kern, A. & Westmoreland, C. (1999). The use of long-term hepatocyte cultures for detecting induction of drug metabolising enzymes: the current status. ECVAM Hepatocytes and Metabolically Competent Systems Task Force Report 1. *ATLA* **27**, 579-605.

Cronin, M.T.D. (2001). Computational methods for the prediction of drug toxicity. *Current Opinion in Drug Discovery and Development* **3**, 292-297.

Dearden, J.C., Barratt, M.D., Benigni, R., Bristol, D.W., Combes, R.D., Cronin, M.T.D., Judson, P.M., Payne, M.P., Richard, A.M., Tichy, M., Worth, A.P., Yourick, J.J. (1997). The development and validation of expert systems for predicting toxicity. The report and recommendations of an ECVAM/ECB workshop (ECVAM workshop 24). *ATLA* **25**, 223-252.

Doehmer, J., Buters, J.T., Luch, A., Soballa, V., Baird, W.M., Morisson, H., Stegeman, J.J., Townsend, A.J., Greenlee, W.F., Glatt, H.R., Seidel, A., Jacob, J. & Greim, H. (1999). Molecular studies on the toxifying effects by genetically engineered cytochromes P450. *Drug Metabolism Reviews* **31**, 423-435.

Eaton, D.L. & Bammler, T.K. (1999) Concise review of the glutathione S-transferases and their significance to toxicology. *Toxicological Sciences* **49**, 156-164.

Friedberg, T., Pritchard, M.P., Bandera, M., Hanlon, S.P., Yao, D., McLaughlin, L.A., Ding, S., Burchell, B. & Wolf, C.R. (1999). Merits and limitations of recombinant models for the study of human P450-mediated drug metabolism and toxicity: an intralaboratory comparison. *Drug Metab Rev* **31**, 523-544.

Glatt H. (2000). Sulfotransferases in the bioactivation of xenobiotics. *Chemico-Biological Interactions* **129,** 141-70.

Gomez-Lechon, M.J., Donato, T., Jover, R., Rodriguez, C., Ponsoda, X., Glaise, D., Castell, J.V. & Guguen-Guillouzo, C. (2001). Expression and induction of a large set of drug-metabolising enzymes by the highly differentiated human hepatoma cell line BC2. *European Journal of Biochemistry* **268**, 1448-1459.

Greene, N., Judson, P.N., Langowski, J.J. & Marchant, C.A. (1999). Knowledge-based expert systems for toxicity and metabolism predictions: DEREK, StAR and METEOR. *SAR & QSAR in Environmental Research* **10**, 299-314.

Hengstler, J.G., Utesch, D., Steinberg, P., Platt, K.L., Diener, B., Ringel, M., Swales, N., Fischer, T., Biefang, K., Gerl, M., Bottger, T. & Oesch, F. (2000). Cryopreserved primary hepatocytes as a constantly available in vitro model for the evaluation of human and animal drug metabolism and enzyme induction. *Drug Metabolism Reviews* **32**, 81-118.

Kassahun, K., Pearson, P.G., Tang, W., McIntosh, I., Leung, K., Elmore, C., Dean, D., Wang, R., Doss, G & Baillie, T.A. (2001). Studies on the metabolism of troglitazone to reactive intermediates *in vitro* and *in vivo*. Evidence for novel biotransformation pathways involving quinone methide formation and thiazolidinedione ring scission. *Chemical Research in Toxicology* **4**, 62-70.

Klopman, G., Diumayuga, M. & Galafgous, J. (1994). META. 1. A program for the evaluation of metabolic transformation of chemicals. *Journal of Chemical Information and Computer Sciences* **34**, 1320-1325.

LeCluyse, E.L. (2001). Human hepatocyte culture systems for the *in vitro* evaluation of cytochrome P450 expression and regulation. *European Journal of Pharmaceutical Sciences* **13**, 343-368.

Lewis, D.F.V., Ioannides, C. & Parke, D.V. (1998). An improved and updated version of the compact procedure for the evaluation of P450-mediated chemical activation. *Drug Metabolism Reviews* **30**, 709-737

Lewis, D.F.V., Lake, B.G., George, S.G., Dickins, M., Eddershaw, P.J., Tarbit, M.H., Beresford, A.P., Goldfarb, P.S. & Guengerich, F.P. (1999). Molecular modelling of CYP1 family enzymes CYP1A1, CYP1A2, CYP1A6, and CYP1B1 based on sequence homology with CYP1O2. *Toxicology* **139**, 53-79.

Li, A.P., Lu C., Brent J.A., Pham C., Fackett A., Ruegg C.E. & Silber P.M. (1999). Cryopreserved human hepatocytes: characterization of drug-metabolizing enzyme activities and applications in higher throughput screening assays for hepatotoxicity, metabolic stability, and drug-drug interaction potential. *Chemico-Biological Interactions* **121,** 17-35.

Linget, J.M. & du Vignaud, P. (1999). Automation of metabolic stability studies in microsomes, cytossol and plasma using a 215 Gilson liquid handler. *Journal of Pharmaceutical and Biomedical Analysis* **19**, 893-901.

Ma, Y., Lang, L., Kiesewetter, D.O., Jagoda, E., Sassaman, M.B, Der, M. & Eckelman, W.C. (2001). Liquid chromatography-tandem mass spectrometry identification of metabolites of two 5-HT1A antagonists, N-[2-[4-(2-methoxylphenyl)piperazino]ethyl]-N-(2-pyridyl) trans- and cis-4-fluorocyclohexane carboxamide, produced by human and rat hepatocytes. *Journal of Chromatography B - Biomedical Applications* **755**, 47-56.

MacGregor, J.T., Collins, J.M., Sugiyama, Y., Tyson, C.A., Dean, J., Smith, L., Andersen, M., Curren, R.D., Houston, J.B., Kadlubar, F.F., Kedderis, G.L., Krishnan, D., Li, A.P., Parchment, R.E., Thummel, K., Tomaszewski, J.E., Ulrich, R., Vickers, A.E.M. & Wrighton, S.A. (2001). In vitro human tissue models in risk assessment: report of a consensus-building workshop. *Toxicological Sciences* **59**, 17-36.

Masimirembwa, C.M., Otter, C., Berg, M., Jonsson, M., Leidvik, B., Jonsson, E., Johansson, T., Backman, A., Edlund, A. & Andersson, T.B. (1999). Heterologous expression and kinetic characterization of human cytochromes P-450: validation of a

pharmaceutical tool for drug metabolism research. *Drug Metabolism and Disposition* **27**, 1117-1122.

Miller, V.P. (2000), Fluorometric high-throughput screening for inhibitors of cytochrome P450. *Annals of the New York Academy of Sciences* **919**, 26-32.

Ozawa, S., Ohta, K., Miyajima, A., Kurebayashi, H., Sunouchi, M., Shimizu, M., Murayama, N., Matsumoto, Y., Fukuoka, M. & Ohno, Y. (2000). Metabolic activation of o-phenylphenol to a major cytotoxic metabolite, phenylhydroquinone: role of human CYP1A2 and rat CYP2C11/CYP2E1. *Xenobiotica* **30**, 1005-1017.

Parke, D.V., Ioannides, C. & Lewis, D.F.V. (1990). The Safety evaluation of drugs and chemicals by the use of computer optimised molecular parametric analysis of chemical toxicity (COMPACT). *ATLA* **18**, 91-102.

Philip, P.A., Ali-Sadat, S., Doehmer, J., Kocarek, T., Akhtar, A., Lu, H. & Chan, K.K. (1999). Use of V79 cells with stably transfected cytochrome P450 cDNAs in studying the metabolism and effects of cytotoxic drugs. *Cancer Chemotherapy and Pharmacology* **43**, 59-67.

Riley, R.J., Parker, A.J., Trigg, S. & Manners, C.N. (2001). Development of a generalized, quantitative physicochemical model of CYP3A4 inhibition for use in early drug discovery. *Pharmaceutical Research* **18**, 652-655.

Rodrigues, A.D. (1999). Integrated cytochrome P450 reaction phenotyping: attempting to bridge the gap between cDNA-expressed cytochromes P450 and native human liver microsomes. *Biochemical Pharmacology* **57**, 465-480.

Swales, N.J. & Utesch, D. (1998). Metabolic activity of fresh and cryopreserved dog hepatocyte suspensions. *Xenobiotica* **28**, 937-948.

Schmalix, W.A., Lang, D., Schneider A., Bocker, R., Greim, H. & Doehmer, J. (1996). Stable expression and coexpression of human cytochrome P450 oxidoreductase and cytochrome P450 1A2 in V79 Chinese hamster cells: sensitivity to quinones and biotransformation of 7-alkoxyresorufins and triazines. *Drug Metabolism and Disposition* **24**, 1314-1319.

Sheweita, S.A. (2001). Drug-metabolizing enzymes: mechanisms and functions. *Current Drug Metaboilsm* **1**, 107-132.

Soucek, P. (1999). Expression of cytochrome P450 2A6 in Escherichia coli: purification, spectral and catalytic characterization, and preparation of polyclonal antibodies. *Archives of Biochemistry and Biophysics* **370**, 190-200.

Talafous, J., Sayre, L.M., Mieyal, J.J. et al. (1994). META. 2. A dictionary model of mammalian xenobiotic metabolism. *Journal of Chemical Information and Computer Sciences* **34**, 1326-1333.

Tanaka, E., Terada, M. & Misawa S. (2000). Cytochrome P450 2E1: its clinical and toxicological role. *Journal of Clinical Pharmacy and Therapeutics* **25**, 165-175.

White, R.E. (2000). High-throughput screening in drug metabolism and pharmacokinetic support of drug discovery. Annual Review of Pharmacology and Toxicology **40**, 133-157.

Wolf, C.R. & Smith, G. (1999). Pharmacogenetics. *British Medical Bulletin* **55**, 366-368.

Chapter 8. Target organ and system toxicity

Spielmann, H., Bochkov, N.P., Costa, L., Gribaldo, L., Guillouzo, A., Heindel, J.J., Karol, M., Parchment, R., Pfaller, W., Peraita, P.P. & Zacharewski, T. (1998). Alternative testing methodologies for organ toxicity. 13th meeting of the Scientific Group on Methodologies for the Safety Evaluation of Chemicals (SGOMSEC). *Environmental Health Perspectives* **106**, Supplement 2, 427-439.

Repeat dose toxicity

Combes, RD. (1999). Summary of discussion: In vitro models for investigation of chronic toxicity and reversibility. *Toxicology in Vitro* **13**, 853-857.

Hartung, T., Prieto, P., Ryan, M.P., Schmuck, G., Sladowski, D., Vericat, J.A., Wendel, A., Wolf, A., Zimmer, J. (2001). Novel advanced *in vitro* methods for long-term toxicity testing. The report and recommendations of ECVAM Workshop 45. *ATLA* **29**, 393-426.

Miles, A.T., Hawksworth, G.M., Beattie, J.H. & Rodilla, V. (2000). Induction, regulation, degradation, and biological significance of mammalian metallothioneins. *Critical Reviews in Biochemistry and Molecular Biology* **35**, 35-70.

Scheers, E.M., Ekwall, B.A. & Dierickx, P.J. (2001). *In vitro* long-term cytotoxicity testing of 27 MEIC chemicals on Hep G2 cells and comparison with acute human toxicity data. *Toxicology in Vitro* **15**, 153–161.

Houy, S. (1998). Primary rat hepatocytes in collagen-sandwich configuration. Characterization and first application as a toxicological *in vitro* tool. PhD thesis. Department of Chemistry, University of Kaiserslautern, Germany.

Hanley, A.B., McBride, J., Oehlschlager, S. & Opara, E. (1999). Session 7: In Vitro models for investigation of chronic toxicity and reversibility. Use of a flow cell bioreactor as a chronic toxicity model system. *Toxicology in Vitro* **13**, 847-851.

Hopfer, U., Jacobberger, J.W., Grunet, D.C., Eckert, R.L., Jat, P.S. & Whisett, J.A. (1996). Immortalization of epithelial cells. *American Journal of Physiology* **270**, C1-C11.

Minuth, W.W., Dermietzel, R., Kloth, S. & Hennerkes, B. (1992). A new method culturing renal cells under permanent superfusion and producing a luminal-basal medium gradient. *Kidney International* **41**, 215-219.

Pfaller, W., Felder, E., Koppelstaetter, C., Abbot, S.E. & Westwick, J. (1999). Epiflow: a new cell culture system utilising continous medium exchange and continuous supply with gases. *ATLA* **27** (special issue), 309.

Schmuck, G. & Schlüter, G. (1996). An *in vitro* model for toxicological investigations of environmental neurotoxins in primary neuronal cell cultures. *Toxicology and Industrial Health* **12**, 683-696.

Nephrotoxicity

Gstraunthaler, G. (1988). Epithelial cells in tissue culture. *Renal Physiology and Biochemistry* **11**, 1-42.

Hawksworth, G.M., Bach, P.H., Nagelkerke, J.F., Dekant, W., Diezi, J.E., Harpur, E., Lock, E.A., MacDonald, C., Morin, P-P., Pfaller, W., Rutten, A.A.J.J.L., Ryan, M.P., Toutain, H.J. & Trevisan, A. (1995). Nephrotoxicity testing *in vitro*. The report and recommendations of ECVAM Workshop 10. *ATLA* **23**, 713 - 727.

Morin, J-P., De Broe, M.E., Pfaller, W. & Schmuck, G. (1997). Nephrotoxicity testing *in vitro*: the current situation. ECVAM Nephrotoxicity Task Force Report 1. *ATLA* **25**, 497-503.

Pfaller, W. & Troppmair, E. (2000). Renal transepithelial resistance (TER) and paracellular permeability (PCP) are reliable endpoints to screen for nephrotoxicity. In *Progress in the Reduction, Refinement and Replacement of Animal Experiments* (ed. M. Balls, A.M. Van Zeller & M.E. Halder). pp. 291-304. Amsterdam: Elsevier.

Steinmassl, D., Pfaller, W., Gstraunthaler, G. & Hoffmann, W. (1995). LLC-PK1 cells as a model for in vitro assessment of proximal tubular nephrotoxicity. *In Vitro Cellular and Developmental Biology* **31**, 94-106.

Bach, P.H., Morin, J-P. & Pfaller, W. (1996). Nephrotoxicity - What we have learned and what we still need to know!. *TEN* **3**, 4-13.

Pfaller, W. & Gstraunthaler, G. (1998). Nephrotoxicity testing *in vitro* - What we know and what we need to know. *Environmental health Perspectives* **106**, 559-569.

Neurotoxicity

Abdulla, E.M. & Campbell, I.C. (1993). *In vitro* tests of neurotoxicity. *Journal of Pharmacological and Toxicological Methods* **29**, 69-75.

Abdulla, E.M., Atterwill, C. & Campbell, I.C. (1995). Workshop on *in vitro* neurotoxicity testing: the obstacles, the way forward. *Clinical and Experimental Pharmacology and Physiology* **22**, 277-280.

Atterwill, C.K., Bruinink, A., Drejer, J., Duarte, E., Abdulla, E.M., Meredith, C., Nicotera, P., Regan, C., Rodríguez-Farré, E., Simpson, M.G., Smith, R., Veronesi, B.,

Vijverberg, H., Walum, E. & Williams D.C. (1994). *In vitro* neurotoxicity testing. The report and recommendations of ECVAM workshop report 3. *ATLA* **22**, 350-362.

Atterwill, C.K. & Purcell W.M. (1999). Brain spheroid and other organotypic culture systems *in vitro* neurotoxicology. In *Neurotoxicology In Vitro* (ed. V.W. Pentreath). pp. 213-238. London: Taylor & Francis.

Blaauboer, B.J., Barratt, M.D. & Houston, J.B. (1999) The integrated use of alternative methods in toxicological risk evalutaion. ECVAM Integrated Testing Strategies Task Force Report 1. *ATLA* **27**, 229-237.

Balls, M. & Walum, E. (1999). Towards the acceptance of *in vitro* neurotoxicity tests. In *Neurotoxicology In Vitro* (ed. V.W. Pentreath). pp. 269-283. London: Taylor & Francis.

Costa, L. (1998). Neurotoxicity testing: a discussion of *in vitro* alternatives. *Environmental Health Perspectives* **106** (Supplement 2), 505-510.

Fielder, R.J., Atterwill, C.K., Anderson, D., Boobis, A.R., Botham, P., Chamberlain, M., Combes, R., Duffy, P., Lewis, R.W., Lumley, C.E., Kimber, I. & Newall, D.R. (1997). BTS working party report on *in vitro* toxicology. *Human and Experimental Toxicol*ogy **16**, S1-S40.

Forsby, A., Pilli, F., Bianchi, V. & Walum, E. (1995). Determination of critical cellular neurotoxic concentrations in human neuroblastoma cell cultures. *ATLA* **23**, 800-811.

Henschler, D., Schmuch, G., van Aerssen, M. & Scchiffmann, D. (1992). The inhibitory effect of neuropathic organophosphate esters on neurite outgrowth in cell cultures: a basis for screening for delayed neurotoxicity. *Toxicology in Vitro* **6**, 327-335.

Honegger, P. & Schilter, B. (1992). Serum-free aggregate cultures of foetal rat brain and liver cells: methodology and some practical applications in neurotoxicology. In *The Brain in Bits and Pieces. In Vitro Techniques in Neurobiology, Neuropharmacology and Neurotoxicology*. (ed. G. Zbinden). pp. 51-79. Zollikon, Switzerland: MTC Verlag.

Kucera, P., Cano, E., Honegger, P., Zijlstra, J.A. & Schmid, B. (1993). Validation of whole chick embryo cultures, whole rat embryo cultures and aggregating embryonic brain cell cultures using six pairs of coded compounds. *Toxicology in Vitro* **7**, 785-798.

Monnet-Tschudi, F., Zurich, M-G. & Honegger, P. (1997). Aggregate cell cultures for neurotoxicity testing: the importance of cell-cell interactions. In *Animal Alternatives, Welfare and Ethics. Proceedings of the Second World Congress on Alternatives and Animal Use in the Life Sciences* (ed. L.F.M. van Zutphen & M. Balls). pp. 641-649. Amsterdam: Elsevier.

OECD (1995a). OECD Guideline for Testing of Chemicals. 418. Delayed Neurotoxicity of Organophosphorous Substances Following Acute Exposure. Paris: Organisation for Economic Cooperation and Development. http://www.oecd.org/ehs/test/testlist.htm.

OECD (1995b). OECD Guideline for Testing of Chemicals. 419. Delayed Neurotoxicity of Organophosphorous Substances: 28-day Repeated Dose Study. Paris: Organisation for Economic Cooperation and Development. http://www.oecd.org/ehs/test/testlist.htm.

OECD (1997). OECD Guideline for Testing of Chemicals. 424. Neurotoxicity Study in Rodents. Paris: Organisation for Economic Cooperation and Development. http://www.oecd.org/ehs/test/testlist.htm.

Pentreath, V.W. (Ed.) (1999). Neurotoxicology *In vitro*. London: Taylor & Francis.

Pentreath, V.W. & Atterwill, C.K. (1996). An *in vitro* neurotoxicity testing scheme: evaluation of cytotoxicity determinations in neural and non-neural cells. *In vitro Toxicology* **9**, 83-92.

Stingele, S., Coecke, S., Nicotera, P. & Balls, M. (1999). EC Patent filed 27th of August 1999 on *Genetically engineered cell lines, and their uses, in particular for neurotoxicity testing*. Patent No. 99402139.2-2106/International No PCT/EP00/08223. Rijwijk, The Netherlands: European Patent Office.

Svendsen, C.N., Borg, M.G., Armstrong, R.J.E., Rosser, A.E., Chandran, S., Ostenfeld, T. & Caldwell, M.A. (1998). A new method for the rapid and long term growth of human neural precursor cells. *Journal of Neuroscience Methods* **85**, 141-152.

Veronesi, B. (1992). *In vitro* screening batteries for neurotoxicants. *Neurotoxicology* **13**, 185-196.

Xie, K. & Harvey, A.L. (1993). Evaluation of nerve cell toxicity *in vitro* by electrophysiological and biochemical methods. *Toxicology in Vitro* **7**, 275-279.

Walum, E., Balls, M., Bianchi, V., Blaauboer, B., Bolcsfoldi, G., Guillouzo, A., Moore, G.A., Odland, L., Reinhardt, C. & Spielmann, H. (1992). ECITTS: an integrated approach to the application of *in vitro* tests systems to the hazard assessment of chemicals. *ATLA* **20**, 406-428.

Weiss, M.T. & Sawyer, T.W. (1993). Cytotoxicity of the MEIC test chemicals in primary neurone cultures. *Toxicology in Vitro* **7**, 653-667.

Williams, S.P., Davenport-Jones, J., Egan, C., O'Hare, S., Cookson, M., McClean, R., Garle, M.J., Pentreathe, V. & Atterwill, C.K. (1994). Phase 1 of an *in vitro* neurotoxicological pre-validation trial. *Toxicology in Vitro* **8**, 799-802.

Williams, S.P., O'Brien, S., Whitmore, K., Purcell, W.M., Cookson, M.R., Mead, C., Pentreath, V.W. & Atterwill, C.K. (1996). An *in vitro* neurotoxicity testing scheme:

evaluation of cytotoxicity determinations in neural and non-neural cells. *In Vitro Toxicology* **9**, 83-92.

Zurich, M.G., Honegger, P., Schilter, B., Costa, L.G. & Monnet-Tschudi, F. (2000). Use of aggregating brain cell cultures to study developmental effects of organophosphorus insecticides. *Neurotoxicology* **21**, 599-605.

Chapter 9. Genotoxicity and carcinogenicity

Aardema, M.J. & Kirsch-Volders, M. (2001). The in vitro micronucleus assay. In *Genetic toxicology and Cancer Risk Assessment* (ed. W.N. Choy), pp. 163–186, Marcel Dekker.

Anon (1997). *Aneuploidy*. ECETOC Monograph No. 27. 172 pp. Brussels: European Centre for Ecotoxicology and Toxicology of Chemicals.

Ashby, J. & Paton, D. (1993). The influence of chemical structure on the extent and sites of carcinogenesis for 522 rodent carcinogens and 55 different human carcinogen exposures. *Mutation Research* **286,** 3-74.

Ashby, J. & Tennant, R.W. (1991). Definitive relationships among chemical structure, carcinogenicity and mutagenicity of 301 chemicals tested by the US NTP. *Mutation Research* **257**, 229-306.

Barratt, M. D. (2000). Principles of toxicity prediction from chemical structure. In *Progress in Reduction, Refinement and Replacement in Animal Experimentation,* (ed. M. Balls, A-M. van Zeller & M.E. Halder). pp. 449-456. Amsterdam: Elsevier Science, BV.

Benigni, R. (2000). Structure-activity relationships in mutagenesis and carcinogenesis. In *Progress in Reduction, Refinement and Replacement in Animal Experimentation,* (ed. M. Balls, A-M. van Zeller & M.E. Halder). pp. 469-478. Amsterdam: Elsevier Science, BV.

Blaauboer, B.J., Barratt, M.D. & Houston, J.B. (1999). The integrated use of alternative methods in toxicological risk evaluation. ECVAM Integrated Testing Strategies Task Force Report 1. *ATLA* **27**, 229-237.

Butterworth, B.E., Connolly, R.B. & Morgan, K.T. (1995). A strategy for establishing mode of action of chemical carcinogens as a guide for approaches to risk assessments. *Cancer Letters* **93**, 129-146.

Choy, W.N. (2001). Genotoxic and non-genotoxic mechanisms of carcinogenesis. In *Genetic toxicology and Cancer Risk Assessment* (ed. W.N. Choy). pp. 47–72. Marcel Dekker.

Combes, R.D. (1992). The *in vivo* relevance of *in vitro* genotoxicity assays incorporating enzyme activation systems. In *Progress in Drug Metabolism Volume 13* (ed. C.G. Gibson). pp. 295-321. London: Taylor and Francis Limited.

Combes, R.D. (1995). Regulatory genotoxicity testing: a critical appraisal. *ATLA.* **23,** 352-379.

Combes, R.D. (1997). Detection of nongenotoxic carcinogens: major barriers to replacement of the rodent bioassay. In *Animal Alternatives, Welfare and Ethics. Proceedings of the Second World Congress on Alternatives and Animal Use in the Life Sciences* (ed. L.F.M. van Zutphen & M. Balls). pp. 627-634. Amsterdam: Elsevier.

Combes, R.D. (1999). Mutagenicity. In *Cosmetics: Controlled Efficacy Studies and Regulation* (ed. P. Elsner, H. Merk & H.I. Mailbach). pp. 291-308. Springer-Verlag.

Combes, R.D. (2000a). Endocrine disruptors – a critical review of *in vitro* and *in vivo* testing strategies for assessing their toxic hazard to humans. *ATLA* **28**, 81-118.

Combes, R.D. (2000b). The use of structure-activity relationships and markers of cell toxicity to detect non-genotoxic carcinogens. *Toxicology In Vitro* **14**, 387-399.

Combes, R.D. & Judson, P. (1995). The use of artificial intelligence systems for predicting toxicity. *Pesticide Science* **45**, 179-194.

Combes, R., Balls, M., Curren, R., Fischbach, M., Fusenig, N., Kirkland, D., Lasne, C., Landolph, J., LeBoeuf, R., Marquardt, H., McCormick, J., Müller, L, Rivedal, E., Sabbioni, E., Tanaka, N., Vasseur, P. & Yamasaki, H. (1999). Cell transformation assays as predictors of human carcinogenicity. *ATLA* **27**, 745-767.

Crespi, C. (1995). Use of genetically engineered cells for genotoxicity testing. In *Environmental Mutagenesis* (ed. D.H. Phillips & S. Venitt). pp. 233-260. Oxford: S.Bios Scientific Publishers.

D'Arcy, P.F. & Harron, D.W.G. (1997). *Proceedings of the Fourth International Conference on Harmonisation, Brussels, 1997*. Belfast: Queen's University.

Dearden, J.C., Barratt, M.D., Benigni, R., Bristol, D.W., Combes, R.D., Cronin, M.T.D., Judson, P.M., Payne, M.P., Richard, A.M., Tichy, M., Worth, A.P., Yourick, J.J. (1997). The development and validation of expert systems for predicting toxicity. The report and recommendations of an ECVAM/ECB workshop (ECVAM workshop 24). *ATLA* **25**, 223-252.

DeVaney, T.T., Tritthart, H.A., Ahammer, H. & Helige, C. (1997). In vitro evaluation of cancer cell invasiveness by confocal laser scan techniques. In *Animal Alternatives, Welfare and Ethics. Proceedings of the Second World Congress on Alternatives and Animal Use in the Life Sciences* (ed. L.F.M. van Zutphen & M. Balls). pp. 329-335. Amsterdam: Elsevier.

DOH (Department of Health) (1989). *Guidelines for the Testing of Chemicals for Mutagenicity*. London: HMSO.

DOH (Department of Health) (2000). *Report on Health and Social Subjects, Guidelines for the Testing of Chemicals for Mutagenicity*. Committee on Mutagenicity of Chemicals in Food, Consumer Products and the Environment. London: HMSO.

Fan, A.M. & Howd, R.A. (2001). Quantitative risk assessment of nongenotoxic carcinogens. In *Genetic toxicology and Cancer Risk Assessment* (ed. W.N. Choy). pp. 299–320. Marcel Dekker.

Fielden, M.R. & Zacharewski, T.R. (2001). Challenges and limitations of gene expression profiling in mechanistic and predictive toxicology. *Toxicological Sciences* **60**, 6-10.

Fielder, R.J., Atterwill, C.K., Anderson, D., Boobis, A.R., Botham, P., Chamberlain, M., Combes, R., Duffy, P., Lewis, R.W., Lumley, C.E., Kimber, I. & Newall, D.R. (1997). BTS working party report on *in vitro* toxicology. *Human and Experimental Toxicology* **16**, S1-S40.

Fry, J. & George, E. (2001). *In vitro* toxicology in the light of new technologies. *ATLA* **29**, 745-748.

Fusenig, N.E. & Boukamp, P. (1998). Multiple stages and genetic alterations in immortalization, malignant transformation, and tumor progression of human skin keratinocytes. *Molecular Carcinogenesis* **23**, 144-158.

Gottmann, E., Kramer, S., Pfahringer, B. & Helma, C. (2001). Data quality in predictive toxicology: reproducibility of rodent carcinogenicity experiments. *Environmental Health Perspectives* **109**, 509-514.

ICH (1997a). Report of Fourth International Conference on Harmonisation, Brussels, July, 1997, Genotoxicity: Standard Battery of Tests. Step 4 Report. ICH.

IEH (2002). Assessment of chemical carcinogens: background to general principles of a weight of evidence approach. 17 pp. Leicester, UK: Medical Research Council Institute for Environment and Health. Available on the internet: http://www.le.ac.uk/ieh/

Kirsch-Volders, M., Sofuni, T., Aardema, M., Albertini, S., Eastmond, D., Fenech, M., Ishidate, M. Jr, Lorge, E., Norppa, H., Surralles, J., von der Hude, W. & Wakata, A. (2000). Report from the In Vitro Micronucleus Assay Working Group. *Environmental and Molecular Mutagenisis* **35**, 167-172.

Klopman, G. & Rosenkranz, H.S. (1994). Approaches to SAR in carcinogenesis and mutagenesis. Prediction of carcinogenicity/mutagenicity using MULTI-CASE. *Mutation Research* **305**, 33-46.

LeBoeuf, R.A., Kerckaert, G.A., Aardema, M.J., Gibson, D.P., Brauninger, R. & Isfort, R.J. (1996). The pH 6.7 Syrian hamster embryo cell transformation assay for assessing the carcinogenic potential of chemicals. *Mutation Research* **356**, 85-127.

LeBoeuf, R.A., Kerckaert, K.A., Aardema, M.J. & Isfort, R J. (1999). Use of Syrian hamster embryo and Balb/c 3T3 cell transformation for assessing the carcinogenicity potential of chemicals. In *The Use of Short- and Medium-term Tests for Carcinogenic Hazard Eualuation. IARC Scientific Publications No. 146* (ed. D.B. McGregor, J.M. Rice & S. Venitt), pp. 409-425. Lyon, France: IARC.

Lewis, D.F. & Langley, G.R. (1996). A validation study of the COMPACT and HazardExpert techniques with 40 chemicals. *Mutation Research* **369**, 157-174.

Lewis, D.F.V., Ioannides, C. & Parke, D.V. (1996). COMPACT and Molecular Structure in Toxicity Assessment: A Prospective Evaluation of 30 Chemicals Currently Being Tested for Rodent Carcinogenicity by the NCI/NTP. *Environmental Health Perspectives* **104S**, 1011-1016.

Long, A. & Combes, R.D. (1995). Using DEREK to predict the activity of some carcinogens/mutagens found in foods. *Toxicology in Vitro* **9,** 563-569.

Macé, K., Offord, E.A. & Pfeifer, A.M.A. (1997). Drug metabolism and carcinogen activation studies with human genetically engineered cells. In *In Vitro Methods in Pharmaceutical Research* (ed. J.V. Castell & M.J. Gomez-Lechon). pp. 433–456. London: Academic Press.

Mareel, M.M., Kint, J. & Meyvisch, C. (1979). Methods of study of the invasion of malignant C3H-mouse fibroblasts into embryonic chick heart in vitro. *Cell Pathology* **30**, 95-111.

Mauthe, R.J., Gibson, D.P., Bunch, R.T. & Custer, L. (2001). The syrian hamster embryo (SHE) cell transformation assay: review of the methods and results. *Toxicologic Pathology* **29**, Supplement 138-146.

Mitchell, I. deG. & Combes, R.D. (1996). *In vitro* genotoxicity and cell transformation assessment. In *In Vitro Methods in Pharmaceutical Research* (ed. J.V. Castell & M.J. Gomez-Lechon). pp. 317-352. London: Academic Press.

Parry, J.M. (1994). Detecting and predicting the activity of rodent carcinogens. *Mutagenesis* **9**, 3-5.

Parry, E.M. & Parry, J.M. (1995). *In vitro* cytogenetics and aueploidy. In *Environmental Mutagenesis* (ed. D.H. Phillips & S. Venitt). pp. 121-139. Oxford: S.Bios Scientific Publishers.

Parry, E.M., Henderson, L. & Mackay, J.M. (1995). Procedures for the detection of chemically induced aneuploidy: recommendations of a UK Environmental Mutagen Society Working Group. *Mutagenesis* **10**, 1-14.

Ridings, J.E., Barratt, M.D., Cary, R., Earnshaw, C.G., Eggington, C.E., Ellis, M.K., Judson, P.N., Langowski, J.J., Marchant, C.A., Payne, M.P., Watson, W.P. & Yih, T.D. (1996). Computer prediction of possible toxic action from chemical structure: an update on the DEREK system. *Toxicology* **106,** 267-279.

Robertson, D.G. et al. (2002). A report of the ILSI study on alternatives to the rodent bioassay for carcinogen identification. *Toxicological Sciences* (in press).

Smolle, J., Helige, C. & Tritthart, H.A. (1992). An image analysis and statistical evaluation program for the assessment of tumor cell invasion in vitro. *Analytical Cellular Pathology* **4**, 49-57.

Stopper, H. & Müller, S.O. (1997). Micronuclei as a biological endpoint for genotoxicity: A minireview. *Toxicology in Vitro* **11,** 661-667.

Tennant, R.W., Stasiewicz, S., Mennear, J., French, J.E. & Spalding, J.W. (1999). Genetically altered mouse models for identifying carcinogens. In *The Use of Short- and Medium- Term Tests for Carcinogens and Data on Genetic Effects in Carcinogenic Hazard Evaluation. IARC Scientific Publications No. 146* (ed. D.B. McGregor, J.M. Rice & S. Venitt). pp. 123–150. Lyon, France: IARC.

Tsuchiya, T. et al. (1999). An interlaboratory validation study of the improved transformation assay employing Balb/c 3T3 cells: results of a collaborative study on the two-stage cell transformation assay by the Non-genotoxic Carcinogen Study Group. *ATLA* **27**, 685–702.

van Zeller, A-M. & Combes, R.D. (1999). Transgenic mouse bioassays for carcinogenicity testing: a step in the right direction? *ATLA* **27**, 839-846.

Wiebel, F.J., Andersson, T.B., Casciano, D.A., Dickins, M., Fischer, V., Glatt, H.., Horbach, J., Langenbach, R.J., Luyten, W., Turchi, G. & Vandewalle, A. (1997). Genetically engineered cell lines: characterisation and applications in toxicity testing. *ATLA* **25**, 625-639.

Zeiger, E. (2001). Genetic toxicity tests for predicting carcinogenicity. In *Genetic Toxicology and Cancer Risk Assessment* (ed. W.N. Choy). pp. 29–46. Marcel Dekker.

Chapter 10. Reproductive toxicity

Allenby, G., Foster, P.M. & Sharpe. R.M. (1991). Evaluation of changes in the secretion of immunoactive inhibin by adult rat seminiferous tubules in vitro as an indicator of early toxicant action on spermatogenesis. *Fundamental and Applied Toxicology* **16**, 710-724.

Audus, K.L. (1999). Controlling drug delivery across the placenta. *European Journal of Pharmaceutical Sciences* **8**, 161-165.

Baker, V. (2001). Endocrine disruptors – testing strategies to assess human hazard. *Toxicology in vitro* **15**, 413-419.

Brown, N.A., Spielmann, H., Bechter, R., Flint, O.P., Freeman, S.J., Jelinek, R.J., Koch, E., Nau, H., Newall, D.R., Plamer, A.K., Renault, M.F., Repetto, M.F., Vogel, R. & Wiger, R. (1995). Screening chemicals for reproductive toxicity: the Current

Alternatives. The report and recommendations of ECVAM workshop 12. *ATLA* **23**, 868-882.

Brun, H.P., Leonard, J.F, Moronvalle, V., Caillaud, J.M., Melcion, C. & Cordier, A. (1991), Pig Leydig cell culture: a useful in vitro test for evaluating the testicular toxicity of compounds. *Toxicology and Applied Pharmacology* **108**, 307-320.

Chen, Y.J., Lin-Chao, S., Huang, T.S., Yan, M.L. & Lu, F.J. (2001). Humic acid induced growth retardation in a sertoli cell line, TM4. *Life Sciences* **69**, 1269-1284

Combes, R.D. (2000). Endocrine disruptors – a critical review of *in vitro* and *in vivo* testing strategies for assessing their toxic hazard to humans. *ATLA* **28**, 81-118.

Cooke, B.A., Choi, M.C., Dirami, G. & Lopez-Ruiz, M. (1992). Control of steroidogenesis in Leydig cells. *Journal of Steroid Biochemistry and Molecular Biology* **43**, 445-449.

Cooke, B.A. (1998). In vitro models for the investigation of reproductive toxicology in the testis. *Advances in Experimental Medicine and Biology* **444**, 95-102.

Cuzin, F. & Rassoulzadegan, M. (1998), Cell culture systems for the analysis of the male germinal differentiation. In *Reproductive Toxicology, In vitro germ cell developmental toxicology: from science to social and industrial demand* (ed. Del Mazo). pp. 41-49 Plenum Press.

De Felici, M. (1998). Isolation and culture of germ cells from the mouse embryo. In *Cell biology: a laboratory handbook, second addition*. pp. 73-85.

De Felici, M., Di Carlo, A., Dolci, S. & Pesce, M. (1998). Experimental *in vitro* approaches to the study of mouse primordial germ cells. *In Testicular function: from gene expression to genetic manipulation* (ed. Stefanini et al). pp 23-39. Springer.

Fielder, R.J., Atterwill, C.K., Anderson, D., Boobis, A.R., Botham, P., Chamberlain, M., Combes, R., Duffy, P., Lewis, R.W., Lumley, C.E., Kimber, I. & Newall, D.R. (1997). BTS working party report on *in vitro* toxicology. *Human and Experimental Toxicol*ogy **16**, S1-S40.

Hardy, K., Wright, C.S., Franks, S. & Winston, R.M. (2000). In vitro maturation of oocytes. *British Medical Bulletin* **56**, 588-602.

Harrison, C.A., Farnworth, P.G., Chan, K.L., Stanton, P.G., Ooi, G.T., Findlay, J.K. & Robertson, D.M. (2001). Identification of specific inhibin A-binding proteins on mouse Leydig (TM3) and sertoli (TM4) cell lines. *Endocrinology* **142**, 1393-1402.

Hinsch, E., Ponce, A.A., Hagele, W., Hedrich, F., Muller-Schlosser, F., Schill, W.B. & Hinsch, K.D. (1997). A new combined in-vitro test model for the identification of substances affecting essential sperm functions. *Human Reproduction* **12**, 2580.

Janecki, A., Jakubowiak, A. & Steinberger, A. (1992). Effect of cadmium chloride on transepithelial electrical resistance of Sertoli cell monolayers in two-compartment cultures - a new model for toxicological investigations of the "blood-testis" barrier in vitro. *Toxicology and Applied Pharmacology* **112**, 51-57.

Klemm, M., Genschow, E., Pohl, I., Barrabas, C., Liebsch, M. & Spielmann, H. (2001). Permanent embryonic mouse germ cell-lines, an in vitro alternative to in vivo germ cell mutagenicity tests. *ALTEX* **18**, 127-130.

Kumi-Diaka, J., Nguyen, V. & Butler, A. (1999). Cytotoxic potential of the phytochemical genistein isoflavone (4',5',7-trihydroxyisoflavone) and certain environmental chemical compounds on testicular cells. *Biology of the Cell* **91**, 515-523.

Nakatsuji, N. & Chuma, S. (2001). Differentiation of mouse primordial germ cells into female or male germ cells. *International Journal of Developmental Biology* **45**, 541-548.

Scholz, G., Pohl, I., Genschow, E., Klemm, M. & Spielmann, H. (1999). Embryotoxicity screening using embryonic stem cells *in vitro*: correlation to *in vivo* teratogenicity. *Cell Tissues Organs* **165**, 203-211.

Shamblott, M.J., Axelman, J., Littlefield, J.W., Blumenthal, P.D., Huggins, G.R., Cui, Y., Cheng, L. & Gearhart, J.D. (2001). Human embryonic germ cell derivatives express a broad range of developmentally distinct markers and proliferate extensively in vitro. *Proceedings of the National Academy of Sciences USA* **98**, 113-118.

Slott, V.L., Suarez, J.D., Poss, P.M., Linder, R.E., Strader, L.F. & Perreault, S.D. (1993). Optimization of the Hamilton-Thorn computerized sperm motility analysis system for use with rat spermatozoa in toxicological studies. *Fundamental and Applied Toxicology* **21**, 298-307.

Skopek. T.R. (1998). Comparison of motility and membrane integrity to assess rat sperm viability. *Reproductive Toxicology* **12**, 105-114.

Spielmann, H. (1998). Reproduction and development. *Environmental Health Perspectives* **106** (Supplement 2), 571-576.

Telfer, E.E. (1998). *In vitro* models for oocyte development. *Theriogenology* **49**, 451-460.

Chapter 11. Endocrine disruption

Anon. (1996a). *Environmental Oestrogens - A Compendium of Test Methods.* ECETOC Document **33**, 27pp.

Anon. (2001a). US Environmental Protection Agency (2001). *Endocrine disruptor screening program.* Available on the internet:
 http://www.epa.gov/scipoly/oscpendo/index.html.

Anon. (2001b). Communication for the Commission to the Council and the European Parliament, *On the Implementation of the Community Strategy for Endocrine Disruptors: a Range of Substances Suspected of Interfering with the Hormone Systems of Humans and Wildlife (COM(1999)706), Commission Communities COM(2001)262 final.*

Ashby, J. (2000). Validation of *in vitro* and *in vivo* methods for assessing endocrine disrupting chemicals. *Toxicologic Pathology* **28**, 432-437.

Ashby, J. & Lefevre, P.A. (2001). Replacement of surgical castration by GnRH inhibition for rat prostate androgen receptor preparations. *Journa of Applied Toxicology* **21**, 353-354.

Ashby, J., Lefevre, P.A., Deghenghi, R. & Wallis, N. (2001). Replacement of surgical castration by GnRH inhibition or Leydig cell ablation in the male rat Hersherger antiandrogen assay. *Regulatory Toxicology and. Pharmacology* **34**, 188-203.

Baker, V. (2001). Endocrine disruptors B testing strategies to assess human hazard. *Toxicology in Vitro* **15**, 413-419.

Blair, R.M., Fang, H., Branham, W.S., Hass, B.S., Dial, S.L., Moland, C.L., Tong, W., Shi, L., Perkins, R. & Sheehan, D.M. (2000). The estrogen receptor relative binding affinities of 188 natural xenochemicals: structural diversity of ligands. *Toxicological Sciences* **54**, 138-153.

Brodie, A., Lu, Q. & Long, B. (1999). Aromatase and its inhibitors. *Journal of Steroid Biochemistry and Molecular Biology* **69**, 205-210.

Combes, R.D. (2000a). Endocrine disruptors: a critical review of *in vitro and in vivo* testing strategies for assessing their toxic hazard to humans. *ATLA* **28**, 81-118.

Combes, R.D. (2000b). The use of structure-activity relationships and markers of cell toxicity to detect non-genotoxic carcinogens. *Toxicology in Vitro* **14**, 387-399.

Combes, R.D. (2002). How much flexibility is possible when validating new *in vivo* and *in vitro* toxicity test methods? An investigation using the OECD validation of the uterotrophic assay as an example. *ATLA* (in preparation).

Crain, D.A., Rooney, A.A., Orlando, E.F. & Guillette, L.J. Jr. (2000). Endocrine disrupting contaminants and hormone dynamics: lessons from wildlife. In: L. Guillette Jr & D.A Crain (Eds.), *Environmental Endocrine Disruptors, and Evolutionary Perspective,* Taylor & Francis, New York, pp. 1-21.

Harvey, P.W., Rush, K.C. & Cockburn, A. Editors (1999). *Endocrine and Hormonal Toxicology*, John Wiley & Sons, Chichester.

Hashimoto, Y., Moriguchi, Y., Oshima, H., Kawaguchi, M., Miyazaki, K. and Nakamura, M. (2001). Measurement of estrogenic activity of chemicals for the development of new dental polymers. *Toxicology in Vitro* **15**, 421-425.

Holmes, P., Humfrey, C. & Scullion, M. (1998). Appraisal of test methods for sex-hormone disrupting chemicals capable of affecting the reproductive process. Document prepared by the Medical Research Council Institute for Environment and Health. 303 pp. Available on the internet: *http://www.oecd.org/ehs/test/monos.htm.*

Hong, H., Tong, W., Fang, H., Shi, L., Xi, Q., Perkins, R., Walker, J.D., Branham, W. and Sheehan, D.M. (2002). Prediction of estrogen receptor binding for 58,000 chemicals using an integrated system of tree-based model with structural alerts. *Enviromental Health Perspective* **110**, 29-36.

Jacobs, M.N. & Lewis, D.V.F. (2002). Steroid hormone receptors and dietary ligands: a selected review. *Proceedings of the Nutrition Society* **61**, 105-122.

Johansson. M.K., Sanderson, J.T. & Lund, B-O. (2002). Effects of 3-MeSO$_2$-DDE and some CYP inhibitors on glucocorticoid steroidogenesis in the H295R human adrenocortical carcinoma cell line. *Toxicology in vitro* (in press).

Kanno, J., Onyon, L., Haseman, J., Fenner-Crisp, P., Ashby, J. & Owens, W. (2001). The OECD program to validate the rat uterotrophic bioassay to screen compounds for *in vivo* estrogenic responses: phase 1. *Enviromental Health Perspectives* **109**, 785-794.

Kavlock, R.J., Daston, G.P., DeRosa, C., Fenner-Crisp, P., Gray, L.E., Kaattari, S., Lucier, G., Luster, M., Mac, M.J., Maczka, C., Miller, R., Moore, J., Rolland, R., Scott, G., Sheehan, D.M., Sinks, T. & Tilson, H.A. (1996). Research needs for the risk assessment of health and environmental effects of endocrine disruptors: a report of the U.S. - EPA - sponsored workshop. *Environmental Health Perspectives* **104**, 715-740.

Kelce, W.R. & Gray, L.E. Jr. (1999). Environmental androgens: *in vitro* and *in vivo* screening mechanisms. *Laboratory Animals*, **28**, 26-32.

Sharp, R.M. (2001). Hormones and testis development and the possible adverse effects of environmental chemicals. *Toxicology Letters* **120**, 221-232.

Shelby, M.D., Newbold, R.R., Tully, D.B., Chae, K. & Davis, V.L. (1996). Assessing environmental chemicals for estrogenicity using a combination of *in vitro* and *in vivo* assays. *Environmental Health Perspectives* **104**, 1296-1300.

Stoker, T.E., Parks, L.G., Gray, L.E. & Cooper, R.L. (2000). Endocrine-disrupting chemicals: prepubertal exposures and effects on sexual maturation and thyroid function in the male rat. A focus on the EDSTAC recommendations. *Critical Reviews in Toxicology* **30**, 197-252.

Warner, M., Nilsson, S & Gustaffson, J-Å (1999). The estrogen receptor family. *Current Opnion in Obstetrics and Gynecology* **11**, 249-254.

Zacharewski, T. (1998). Identification and assessment of endocrine disruptors: limitations of *in vivo* and *in vitro* assays. *Environmental Health Perspectives* **106 S2**, 577-582.