

# An Overall Strategy for the Testing of Chemicals for Human Hazard and Risk Assessment under the EU REACH System

Robert Combes,<sup>1</sup> Martin Barratt<sup>2</sup> and Michael Balls<sup>1</sup>

<sup>1</sup>FRAME, Russell & Burch House, 96–98 North Sherwood Street, Nottingham, NG1 4EE, UK; <sup>2</sup>Marlin Consultancy, 10 Beeby Way, Carlton, Bedford MK43 7LW, UK

**Summary** — In its White Paper, *Strategy for a Future Chemicals Policy*, published in 2001, the European Commission (EC) proposed the REACH (Registration, Evaluation and Authorisation of CHemicals) system to deal with both existing and new chemical substances. This system is based on a top-down approach to toxicity testing, in which the degree of toxicity information required is dictated primarily by production volume (tonnage). If testing is to be based on traditional methods, very large numbers of laboratory animals could be needed in response to the REACH system, causing ethical, scientific and logistical problems that would be incompatible with the time-schedule envisaged for testing. The EC has emphasised the need to minimise animal use, but has failed to produce a comprehensive strategy for doing so. The present document provides an overall scheme for predictive toxicity testing, whereby the non-animal methods identified and discussed in a recent and comprehensive ECVAM document, could be used in a tiered approach to provide a rapid and scientifically justified basis for the risk assessment of chemicals for their toxic effects in humans. The scheme starts with a preliminary risk assessment process (involving available information on hazard and exposure), followed by testing, based on physicochemical properties and (Q)SAR approaches. (Q)SAR analyses are used in conjunction with expert system and biokinetic modelling, and information on metabolism and identification of the principal metabolites in humans. The resulting information is then combined with production levels and patterns of use to assess potential human exposure. The nature and extent of any further testing should be based strictly on the need to fill essential information gaps in order to generate adequate risk assessments, and should rely on non-animal methods, as far as possible. The scheme also includes a feedback loop, so that new information is used to improve the predictivity of computational expert systems. Several recommendations are made, the most important of which is that the European Union (EU) should actively promote the improvement and validation of (Q)SAR models and expert systems, and computer-based methods for biokinetic modelling, since these offer the most realistic and most economical solution to the need to test large numbers of chemicals.

**Key words:** *biokinetic modelling, decision-trees, ECVAM, EU chemicals policy, non-animal testing scheme, (Q)SAR and expert systems, read-across, risk assessment, threshold of regulation.*

## Introduction

In February 2001, the European Commission (EC) issued a White Paper entitled *Strategy for a Future Chemicals Policy* (1), which was subsequently endorsed by the Member States of the European Union (EU). The legislation and action plan needed for the introduction of the new policy are currently under discussion.

The principal objective of the White Paper is to ensure a high level of protection for human health and the environment, while ensuring the efficient functioning of the chemicals market and safeguarding the innovation and competitiveness of the European chemicals industry. The White Paper is based on the precautionary principle, whereby chemicals are considered to be unacceptably hazardous until proven otherwise, and the substitution of dangerous substances by less dangerous ones is encouraged, wherever possible. The White Paper also states that any new legislation should be based on sound science and also on the established principles of risk assessment.

*Existing chemicals* (i.e. chemicals that were already on the market before new EU legislation on chemicals came into force 18 September 1981) have been registered under a different scheme from that adopted for *new chemicals* (i.e. chemicals marketed after that date). Since the earlier regime was less rigorous in its testing requirements, there is concern that a substantial number of the existing chemicals which are currently marketed, may have been inadequately tested and could therefore be harmful, because of so-called data gaps in the information relating to their hazard potential.

To address this problem, the EC White Paper proposed the establishment of a new system called REACH (Registration, Evaluation and Authorisation of CHemicals), to deal with both existing and new chemical substances. The REACH system is based on a top-down approach to toxicity testing, in which the type of information required is dictated primarily by production volume (tonnage). This “volume-triggered notification system” approach is based on the traditional assumption that the higher the level of production of a substance, the greater the

potential level of overall human exposure. We consider this approach to be fundamentally flawed, for reasons explained below.

The actual tests that will be required to establish hazard for various toxicity endpoints have yet to be determined.

The effects that the implementation of the White Paper could have by increasing the scale of toxicity testing have been the subject of much heated debate, and there have been several predictions that very large numbers of laboratory animals would be needed, and that the demands of any new legislation would far surpass the current capacity of the contract testing industry. Moreover, traditional approaches to toxicity testing would be incompatible with the time-schedule envisaged for compliance with the REACH system. Despite the fact that the EC has emphasised the need to minimise the use of laboratory animals, by maximising the implementation of alternatives and avoiding duplicate testing, it has failed to produce a comprehensive strategy for how this could be achieved, while still meeting the objectives of the White Paper (2).

## Objectives of This Paper

These developments have prompted various organisations to promote the use of new approaches to toxicity testing and risk assessment, with an emphasis on the adoption of integrated testing schemes making maximum use of non-animal testing methods, and a more scientific assessment of human risk from exposure to chemicals.

The European Centre for the Validation of Alternative Methods (ECVAM), in conjunction with a group of external partners, has produced a comprehensive report on the status of non-animal methods for the toxicity testing of chemicals (3). This report indicates how these methods can be used in tier-testing schemes for different endpoints, and also identifies what needs to be done to develop and validate new test methods for endpoints that currently lack acceptable test protocols for regulatory application. The ECVAM document and the strategy proposed in this paper are both focused on human health aspects of chemical risk assessment, although we recognise that environmental and wildlife issues are also addressed by the REACH system.

Our paper proposes an overall scheme whereby the non-animal methods identified and discussed in the ECVAM document could be used in conjunction with each other in a tiered approach, in order to provide a rapid and scientifically justifiable basis for the risk assessment of chemicals for their toxic effects in humans. The scheme is designed to be used in conjunction with the individual testing strategies suggested in the ECVAM report (3). The scheme does not attempt to take account of testing

for the effects of chemicals on the environment and on wildlife, since this would require different considerations, and a separate and different scheme.

The scheme starts with a preliminary risk assessment process, involving the use of available information on exposure and hazard, followed by some relatively simple testing, based on physicochemical properties and (quantitative) structure-activity relationship [(Q)SAR] approaches. (Q)SAR studies involve the generation of equations relating physicochemical properties of molecules to their biological activities, to predict the activities of unknown but related substances. The (Q)SAR approach is used in conjunction with expert system modelling, and consideration of information on metabolism and the identification of the principal human metabolites (to estimate bioavailability).

This information should then be combined with an assessment of potential exposure based on production levels, distribution, and modes of use, to provide for an acceptable and adequate risk assessment. Significant and crucial data gaps, including information on likely human exposure levels, should also be identified.

It is recommended that additional testing should only be required where essential information is missing, rather than testing to cover all data gaps according to a generalised, check-list approach. In other words, the official waiving of the need for hazard data by regulatory agencies, on a case-by-case basis, is an integral component of the scheme.

The selection of tests to provide this essential information should be based on the use of non-animal methods, as far as possible. A further risk assessment is then undertaken, and the scheme also includes a feedback loop, so that all the available and relevant information of sufficiently high quality is used to improve the predictivity of computational expert system methods.

The essential elements of the proposed tier-testing scheme are presented in Table 1, and the overall scheme is outlined in Figure 1. Table 2 lists the available non-animal tests, test batteries and models discussed in the ECVAM report (3), which also discusses their advantages, limitations, validation status, and their deployment as screens and batteries.

## Rationale for the Proposed Scheme

### The concept of a hazard-free environment

One of the basic principles of toxicology is that no substance or product can be regarded as completely non-hazardous. Toxicity is crucially dependent on the biological system, the dose, and the nature of the exposure involved. Thus, it follows that it is impossible to assume absolute safety, and attempts, such as

**Table 1: The essential elements involved in a tier-testing scheme**

---

**Stage 1: Prediction of human exposure**

- Physicochemical properties (to determine bioavailability)
  - Current production levels
  - Mode of production
  - Numbers and types of workers involved in production
  - Distribution and transport
  - Major uses (nature and extent)
  - Previous information on human exposure (workplace, population)
  - Methods of disposal
  - Biodegradability (recalcitrance)
  - Residual environmental levels
  - Clustering of chemicals (according to the above information)
- 

**Stage 2: Preliminary risk assessment and compound prioritisation**

- Combined use of ECETOC (30) and CONCAWE (31) schemes for estimating exposure levels
  - Separate calculations for occupational use (primary, downstream users) and use by general population
  - Prediction of human metabolism and principal human metabolites by using battery of genetically engineered mammalian cell lines expressing human CYP isozymes and phase II metabolising enzymes, and expert system prediction methods
  - Use of all available hazard data (published, unpublished)
  - Use of (Q)SAR & expert system computational predictions of toxicity, taking account of information on metabolism
  - Use of read-across, if possible
  - Use of reverse risk assessment (Figure 2)
  - Use of all information (including whether chemical is a known or suspected carcinogen, mutagen or is toxic to the reproductive system [CMR] or a persistent organic pollutant [POP]) to either: a) classify or b) prioritise for further consideration
- 

**Stage 3: Decision on missing data required to make/improve preliminary risk assessment**

- Use of all previous information to identify data gaps
  - Decision on data required and data not-required, according to information in stages 1 and 2
  - Consider use of alternative products for which risk assessment is satisfactory, or for which known hazard data are not unfavourable
  - If possible, performance of a risk assessment on an alternative product
  - Decision on toxicity endpoints needed to provide the minimum information for making a relevant risk assessment
  - Decision on need for testing (non-animal as far as possible; Table 2)
- 

**Stage 4: Obtaining missing hazard information**

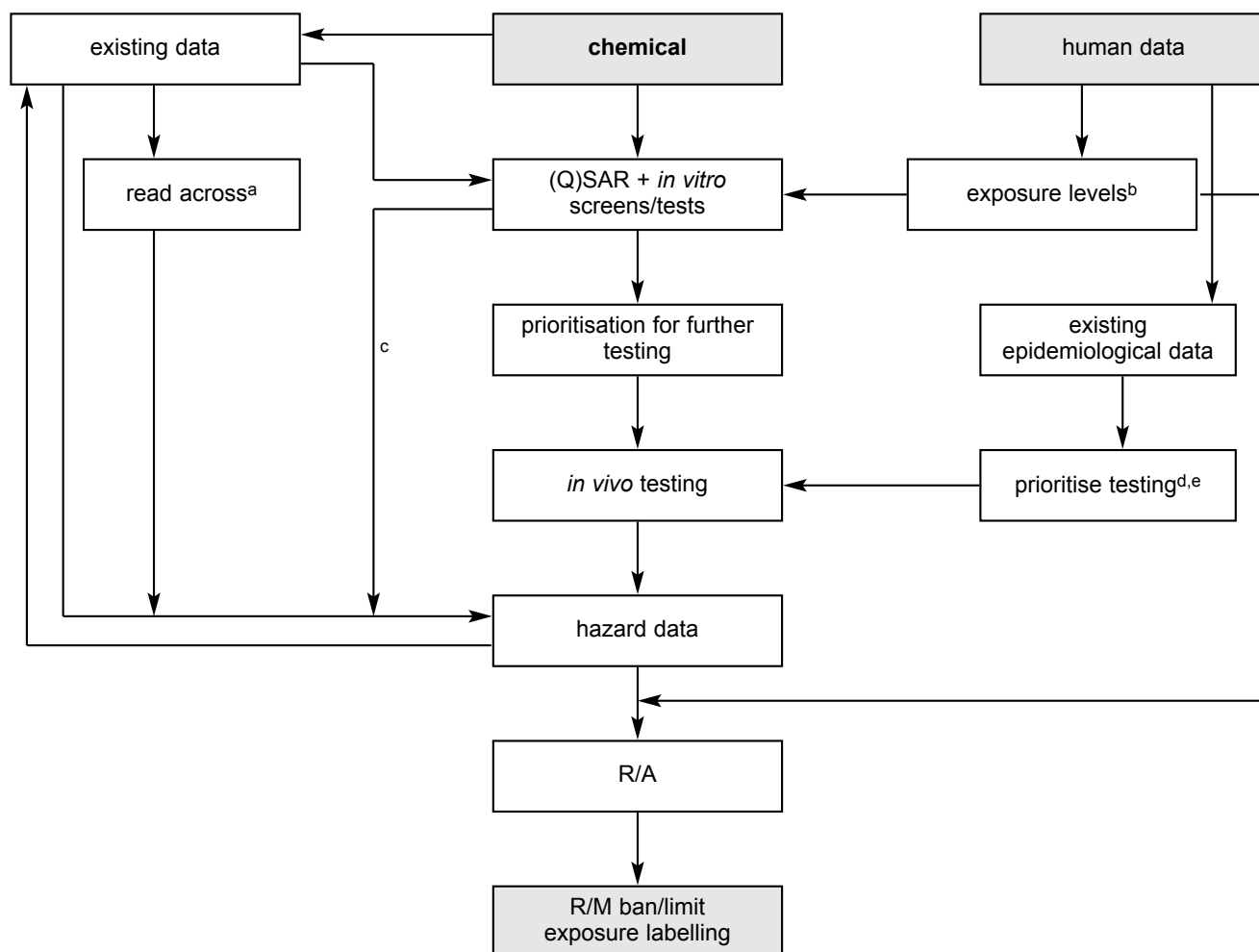
- Assessment of existing information, and encouragement of data sharing
  - Use non-animal approaches as far as possible in an integrated testing scheme (Table 2; Figure 3)
  - Assessment of quality of existing and new hazard data to improve existing (Q)SARs and expert systems rules and to generate new ones (Figure 2)
  - Use of a toxic hazard decision-tree where chemical structure and acute oral lethal toxicity are available (see reference 4)
  - Step-wise application of *in silico*, *in vitro* and *in vivo* methods to obtain new data
  - Performance of risk assessments (by using both *in vitro* and *in vivo* hazard data)
- 

those embodied in the EC White Paper, to achieve an environment completely free from chemical hazard must be judged to be completely misguided.

Toxicologists and regulators define threshold dose-levels (where this can be done), and assess risk as a probability in relation to knowledge of hazard and extent of exposure. Risk management is aimed at minimising the risk, and the only way

in which risk could be avoided altogether would be to ban all chemicals, which of course would be impossible. In fact, many of the chemicals and products in existence are only useful when they are toxic under the variety of conditions prescribed for their use.

However, there is a general lack of public awareness of these issues, as has been all too apparent in

**Figure 1: A scheme for the integrated testing of existing chemicals**

<sup>a</sup>As performed by the Health and Safety Executive; <sup>b</sup>if unavailable, generate before *in vivo* testing; <sup>c</sup>in a few cases, where *in vitro* tests have been accepted as replacements; <sup>d</sup>focus on subchronic/chronic, if acute human data are available;

<sup>e</sup>Prioritisation of testing: low priority and high priority are relative phrases, meaning that chemicals of high priority need to be subjected to risk assessment before chemicals of low priority. Other information, including uses and benefits, availability of substitutes, relative toxic potencies and levels of anticipated human exposures, are all factors that should be taken into account when assigning further priorities within these two broad categories.

the wake of recent problems and controversies, such as those involving BSE, foot and mouth disease, food poisoning, and genetically modified (GM) foods. This is despite the fact that all of us continually weigh up risks and benefits in many aspect of our daily lives, from betting to sunbathing and travelling. There is an important need for those responsible for the EU Chemicals Policy to address this problem and to embark on a policy of public education.

### The need for flexibility

It should usually be unnecessary for a full package of *in vivo* toxicity tests to be undertaken by noti-

fiers, where: a) human exposure is known or predicted to be minimal or non-existent under normal conditions of manufacture and use; b) human exposure is limited by packaging and normal handling, and risk can be virtually eliminated by taking appropriate precautions; c) adequate toxicity data of a suitable quality already exist; d) it is possible to accurately predict the likely toxicity to humans by using knowledge of the physicochemical properties and known biological activities of analogous chemicals with similar functions and uses; e) previous knowledge of human exposure has not revealed any evidence of adverse health effects; f) there are available non-animal replacement tests for particular

toxicity endpoints; g) a toxic chemical can be replaced by a safer alternative; h) a recalcitrant chemical can be replaced by a less, but similarly, effective and useful, less-persistent chemical; and i) the intended use of a chemical is absolutely dependent upon its toxicity to biological systems, in particular or in general, and where direct exposure can reasonably be avoided.

### Situations of low or non-existent human exposure

There will be many cases where it is known that humans are not exposed to particular existing chemical substances, at least under normal conditions of use. There might be concern about accidental exposure, but the likelihood of this could be estimated and included in the risk assessment. This sort of information should be available from details of manufacture and use, and also from any reports related to adverse effects.

### The use of tonnage-triggers for testing

Exposure to a substance does not necessarily imply that it will enter the body, since it might have very low bioavailability, due to its physicochemical properties. Basing a testing strategy on the premise that more exposure occurs with high production/marketing volumes of a very diverse group of substances, irrespective of their physicochemical properties and uses, is therefore fundamentally flawed. It would be much more sensible to cluster the existing chemicals for relatedness, in terms of uses and structural properties, rather than classifying them according to tonnages. There could then be a consideration of known and likely human and environmental exposure levels, as well as of pre-existing information. Bioavailability to different biological systems (such as passage across internal and external epithelial barriers and internal blood-barrier systems) could be predicted from parameters such as octanol/water partition coefficient and molecular size. All this information would be useful for prioritising substances for further testing or registration.

### Other factors affecting human exposure

The nature and extent of human exposure can vary according to the manner in which a substance is packaged and handled. By taking suitable precautions, it is possible to limit, if not eliminate, exposure of the general public under normal and intended circumstances. However, the ease with which this can be achieved will be dic-

**Table 2: Available non-animal approaches for toxicity testing**

---

#### a) Tiered testing schemes/batteries

Physicochemical properties  
Acute toxicity testing  
Corrosivity  
Sensitisation  
Carcinogenicity testing  
Xenobiotic metabolism  
Neurotoxicity  
Genotoxicity  
Hormonal disruption

---

#### b) (Q)SAR and cell culture models

Eye irritation  
Skin irritation  
Phototoxic potential  
Nephrotoxicity  
Reproductive toxicity  
Skin penetration  
Chronic toxicity  
Blood-brain barrier  
Gastrointestinal barrier

---

tated by a variety of factors, including the physical nature of the substance and/or the products which contain it. For example, liquids, dusts, powders, vapours and gases differ according to how they may enter the body. Thus, knowledge of likely form of exposure is essential for deciding the extent and type of toxicity testing required. Moreover, if exposure is limited to occupational activities (for example, manufacturing, packaging, or commercial applications, such as pesticide usage), this should dictate the way in which chemicals are tested, taking into account likely routes of exposure and the likely target organs which might be affected.

If exposure arises solely due to occupational activities, the number of individuals potentially at risk is reduced, not only due to the smaller population size to be considered, but also because the use of protective clothing and other handling precautions, as well as of rigorous disposal procedures, are more practicable. Methods are available for estimating endogenous exposure levels by undertaking the monitoring of individuals for biomarkers of exposure (4).

From an industry perspective, it is important that chemical substances and products are correctly labelled and packaged according to their potential hazard. This applies as much to underestimation of toxic hazard as to its over-estimation. This is because labelling and packaging have significant economic effects by determining the type of transport which is acceptable, and the ease with which a product can be stored and used. Such economic considerations can affect the

extent and nature of any testing which is required, because more-extensive animal testing is often necessary to establish lack of toxicity or low hazard, whereas the results from non-animal screens can be more-easily used for conservative labelling purposes.

### **The use of pre-existing toxicity information**

The EC White Paper emphasises the need to use pre-existing toxicity data and related information to reduce the necessity for new testing, for both logistical and animal welfare reasons. There is also mention of existing EU Directives which encourage data sharing, but their application may be frustrated due to reasons of ownership and confidentiality. Nevertheless, the use of pre-existing information is not given sufficient priority in the EC White Paper, and it is not yet clear how this will be integrated into the overall EU testing strategy.

A wide variety of industrial and other data are publicly available, although a large proportion of this information has probably been obtained by using old protocols, and not according to current Good Laboratory Practice (GLP) standards. However, the industrial companies involved should be encouraged to submit such data to the Competent Authorities, so that they could be used within limits, especially for purposes of prioritising for further testing, and possibly as the basis of initial (Q)SAR models that could be refined later. Maximum use should be made of the information and experience that is being obtained from the ongoing initiative by Human and Environmental Risk Assessment (HERA), comprising a collaboration of the chemical suppliers and manufacturers for the detergent industry, with the objective of providing a common risk assessment framework for these products. Further details of this initiative can be found at [www.Heraproject.com](http://www.Heraproject.com). This is essentially a collaboration between CEFIC (European Chemical Industry Council) and AISE (International Association for Soaps, Detergents and Maintenance Products; i.e. suppliers, formulators and downstream users), all of whom will be required to provide individual risk assessments for chemicals and products under the REACH system. The information from this initiative will be useful for showing how data sharing is being implemented, and will also highlight the kinds of issues that will have to be faced by industry in response to the new legislation.

It is suggested that the EU should require the chemical industry and downstream users to make available all existing data on a substance before any new testing of the substance is permitted. These data should be incorporated into a central database, together with any other relevant information, published or unpublished, and made available to the

Competent Authorities on behalf of notifiers. It is most important that there is an indication of the quality of all such information, including its extent, the endpoints involved, the methods used, and whether or not it was obtained according to GLP. Decisions can then be made as to the extent of use that can be made of the data, and whether or not specific tests need to be repeated or undertaken for the first time.

The proposed EU strategy for existing chemicals appears to ignore the fact that such chemicals differ from new ones in that there will usually be a long history of human exposure to them under conditions of manufacture and use, including under occupational conditions. It is more than likely that the hazard potential of many existing chemicals should already have been detected in terms of adverse effects. This assumption applies mainly to short-term, acute effects, and possibly sub-acute effects. However, it will be unlikely to apply to chronic effects, especially those such as cancer induction, which require long latency periods, and also some reproductive and hereditary effects (for example, germ-cell mutations) that may only be manifested in future generations.

Therefore, it is further suggested that the EU strategy should focus on assessing long-term, rather than acute, toxic effects, and should include investigations of information on all available human adverse effects related to acute and sub-acute toxic hazard. To some extent, the White Paper takes into account the importance of long-term effects, by placing greater emphasis on CMRs (chemicals with the potential to cause mutagenic, carcinogenic and reproductive effects). However, we are proposing that the need to actually generate animal data from acute and sub-acute exposure should be considered in relation to the possibility that human data already exist which are relevant to such exposures. Where such information is lacking, it might be possible to undertake some non-invasive biomonitoring in exposed groups of individuals, to assess the production of biomarkers of effect (4), before proceeding to employ other ways of toxicity testing.

### **The concept of decision-trees**

Various step-wise or decision-tree approaches have been devised to avoid the problem of testing large numbers of chemicals for toxicity, by maximising the use of existing information to predict the likely effects of other chemicals. Decision-trees are particularly important where the demand for data outstrips the supply of the necessary hazard information, a situation which exactly parallels that represented by the resolution to revise the procedures for testing existing and new chemical substances in the EU.

One early example of a decision-tree approach to toxicity is that developed by Cramer *et al.* as long ago as 1978 (5). These authors sought to classify every structurally defined organic and metallo-organic chemical according to criteria largely based on structure and/or on “widely known facts of biochemistry and physiological chemistry”. By using their decision tree, it is possible to place chemicals into one of three categories, representing differing orders of toxic hazard, in order to prioritise them for further testing. Class I substances are those with low toxicity, that should be given a low priority for testing, when combined with evidence for low human exposure. Class III substances are those where it is considered that low toxicity cannot be assumed, or where toxicity is likely. These substances should have a high priority for testing, based on likely levels of human exposure. Class II substances are intermediate between classes I and III, being less harmful than chemicals in Class III, because they lack structural features suggestive of toxicity, although there is no direct evidence to suggest that they are either non-toxic or toxic.

One drawback in applying this particular scheme to chemicals in general is that it was developed for food additives testing, and hence is based only on the use of information on oral toxicity, with chemicals listed in increasing order of No Observed Adverse Effect Level (NOAEL) after oral dosing. There has, however, been interest in the pharmaceutical industry in developing similar schemes to prioritise drug candidate molecules for further investigation (6). This has been provoked by the need to develop rapid screening methods for large numbers of chemicals.

### The threshold of regulatory concern concept

The threshold of regulation concept is based on the premise, originally suggested by Frawley (7), that existing, historical toxicity data for a wide range of substances could be used to formulate a safe maximum level (threshold) of exposure which could be applied generally to any new chemical. If the maximum amounts of the novel chemical in the environment, and in foods, in particular, were known to be, or were expected to be, less than this threshold, it might be possible to exempt the chemical from legislation in the form of petition requirements (8–10).

This concept contrasts with the traditionally accepted view that setting a generalised cut-off level for regulation is unjustified, and that any threshold can only be applied on a case-by-case basis. This well-established approach involves the use of known or predicted NOAEL values for human exposure, which are then used in conjunction with uncertainty factors to set exposure limits for specific chemicals (for example, Average Daily Intake [ADI] values for food additives).

The threshold of regulatory concern concept is also consistent with decision-tree approaches to the prediction of toxicity, as discussed above.

The original proposal for the threshold of regulation concept was designed to deal with the problem of so-called indirect food additives that are present in foods, often in very low amounts, by virtue of their migration from packaging materials, or due to their presence as contaminants as a result of the use of pesticides. There is, however, no *a priori* reason why the concept could not be applied to any chemical or product.

Nevertheless, there has been considerable debate about whether the concept is valid and how much it could be extended beyond its use with indirect food additives. There are two principal problems: a) defining what the exact value of the threshold should be; and b) the extent to which thresholds apply to different mechanisms of toxicity.

With regard to the latter question, it is generally considered that chemicals that act as carcinogens via a genotoxic mechanism (by directly eliciting DNA damage), and which can also be mutagenic, lack threshold dose-levels. This is because it is considered in theory that a single molecule of a chemical can result in a mutational change in DNA, which can ultimately result in tumorigenesis (according to the one-hit theory of carcinogenesis).

It is for this reason that we recommend that any integrated testing scheme incorporating this approach should require that the genotoxicity of the chemical should be determined first. Initially, this should be based on *in vitro* testing, although there might be a need to confirm lack of genotoxicity by undertaking some *in vivo* testing, especially where there was concern that the chemical could be uniquely positive in whole-animal systems. If the chemical proved to be non-genotoxic, consideration should be given to applying the threshold of regulation concept. However, before this practice becomes widespread, more research is needed to investigate its general applicability. Nevertheless, it should be noted that the International Life Science Institute (ILSI) Europe is to recommend a step-wise approach for the application of the threshold of regulation concept for food safety assessment (11).

### Read-across

Read-across (12) is an approach whereby inferences are made about the likely toxicity of a chemical from its similarity to another chemical with a known toxicity profile, with respect to physico-chemical properties, manufacturing processes, uses, and likely human exposure.

Barratt (13) gives hypothetical examples of how this approach could be used to evaluate the systemic toxicities of a series of alkyl betaine surfactants.

Specific *in vitro* tests could be used to provide useful information for read-across approaches to be applied in order to extrapolate from chemicals with *in vivo* data to chemicals that are structurally related, but which have not been tested in animals.

The use of read-across between related chemicals appears to be acceptable to some regulatory authorities, such as the Health & Safety Executive (HSE) in the UK (12) and the Federal German Authorities (14), provided that a valid scientific basis can be provided to support it.

However, it is important to realise that this approach has to be implemented on a case-by-case basis, and that its more widespread use will depend on its further endorsement.

### Reverse risk assessment

When developing a new chemical, it is usual for a standard set of hazard assessment tests to be performed, at the end of which a risk assessment is carried out in the context of the intended use of the new chemical. Past experience has shown that, in many cases, the margin of safety for the prospective use (the safety factor, or the magnitude by which

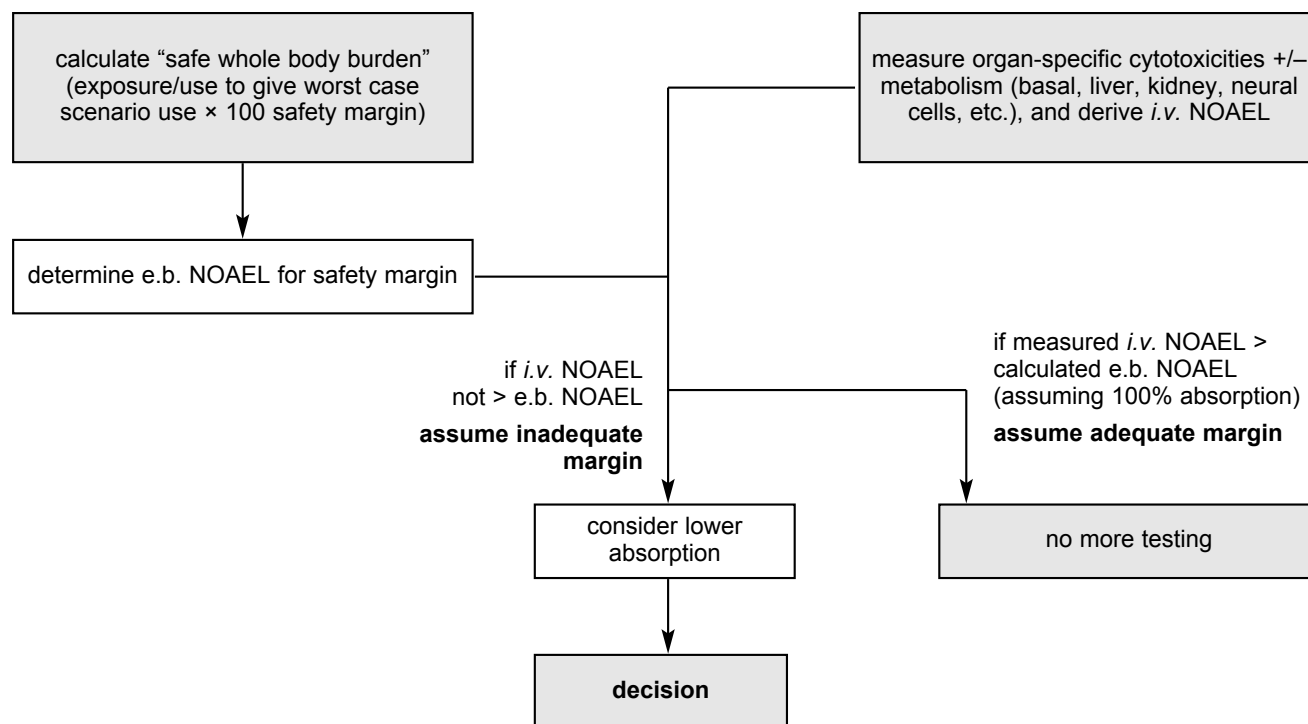
the NOAEL exceeds the known or estimated exposure level) can be found to be so high that it becomes clear with hindsight that much of the testing that had already been undertaken, particularly that at low dose levels, had been completely unnecessary.

It is possible to avoid such superfluous testing by applying the reverse risk assessment process. This entails selecting testing requirements specifically with regard to likely expected exposure to the new chemical. In some cases, where the prospective exposure of the new chemical is extremely low, it might be deemed unnecessary for any animal tests to be performed at all.

The reverse risk assessment process is outlined in Figure 2. The first step is to perform an exposure assessment based on experimental measurements and/or predictions. This should be a worst-case scenario, assuming 100% absorption. An exposure-based NOAEL level is then determined, which incorporates a defined safety factor (conventionally, a factor of 100 is used). An *in vitro* NOAEL is then determined, by measuring organ-specific cytotoxicities with cultured mammalian cells.

If the *in vitro* NOAEL is higher than the exposure-based NOAEL, this means that the chemical has a lower toxicity than would cause concern for

**Figure 2: A scheme for the reverse risk assessment of chemicals**



NOAEL is the No Observed Adverse Effect Level; e.b. = exposure-based; i.v. = *in vitro*.

the likely exposure. For example, an *in vitro* NOAEL 100 times higher than the NOAEL would lead to a final safety factor of 10,000.

If the *in vitro* NOAEL is of the same order or slightly lower than the NOAEL calculated from the exposure, the possibility that absorption might be less than 100% should be considered, before a decision on further testing is taken.

If the *in vitro* NOAEL is less than the exposure-based NOAEL, a decision can be made, either that the margin of safety is inadequate for the proposed exposure, or that further testing may be required before a final decision is taken, depending on the magnitude of the difference between the two values.

The reverse risk assessment approach is used in-house by a number of companies before embarking on a programme of toxicity testing.

### Predicting metabolism

Metabolism is a key determinant of toxicity, since a large number of chemicals are only toxic due to the generation of reactive intermediates, while others are rapidly detoxified. There is therefore a need for predicting the susceptibility of a chemical to metabolism, and for identifying the principal metabolites likely to be generated with respect to the conditions of exposure, target species and target organs involved.

Metabolism is also one of the principal factors which influence the rate and route of elimination of a chemical from the body, since it can change the physicochemical properties of a chemical, and thus its transport and partitioning in biological systems.

The pattern of metabolism to which a compound is susceptible provides useful information on its likely toxicity in any given biological system. Thus, if the chemical is subjected to phase I metabolism by the mixed function oxidases of the cytochrome P450 (CYP) super-gene family of isoenzymes, it is likely to be converted to reactive intermediates. Many genetic polymorphisms in CYP isozyme distribution exist, as well as differences in the profiles of CYP isozymes in different species. Also, in the same species, particular CYP isozymes are associated with the metabolism and activation of specific chemical groups. For example, CYP1A1 is very active in metabolising polycyclic aromatic hydrocarbons, CYP2B1 and CYP1A2 preferentially metabolise aromatic amines, and CYP2E1 metabolises low molecular weight chemicals, whereas CYP3A4 metabolises large ones. CYP2D6 is very important in human drug metabolism.

On the other hand, if a chemical is primarily subjected to phase II metabolism, this will often lead to detoxification via a series of potential conjugation reactions (for example, glucuronidation or glutathione binding).

Determining the susceptibility of a chemical to metabolism, and establishing the nature and relative

quantities of its main (human) metabolites, are therefore indispensable processes in the overall chemical hazard assessment process. Moreover, this is essential, in order to maximise the utility of any computer-based prediction systems. Identifying metabolites can be time-consuming, as it relies on the use of radio-labelled chemicals and sophisticated analytical techniques, such as Nuclear Magnetic Resonance, High Performance Liquid Chromatography and antibody binding. However, it is possible to undertake much useful work *in vitro*, by using pure enzymes, subcellular metabolising fractions (such as a mitochondrial supernatant [S9] fraction, microsomes or cytosol), or metabolically-competent hepatocytes, which can also be co-cultured with toxicity indicator cells (15). The source of metabolism can be varied according to sex, species and tissue, as well as enzyme, or by the use of selective enzyme inducers or inhibitors specific for particular isozymes.

However, a better approach is to use immortalised cells or established cell lines genetically engineered to express various phase I and phase II enzymes, either singly or in combination. Such genetically modified cell lines can be used in test batteries to detect the cytotoxic, cytogenetic and mutagenic effects of chemicals, and to characterise whether they can be activated or detoxified by specific human enzymes, in order to identify their principal metabolites (16, 17). Several genetically engineered cell lines expressing differing forms of the human CYP2D6 isozyme have been used to investigate the phenomenon of genetic polymorphism, in which humans vary in their capacities to metabolise more than 50 drugs that are known to be substrates of the isozyme (18).

In the above way, it is currently possible for one laboratory to screen some 200 chemicals per month (Johannes Doehmer, personal communication). It is therefore recommended that the use of batteries of cell lines with different and specific metabolising capacities for screening large numbers of chemicals, should be actively developed and promoted, especially as part of an integrated testing strategy for existing chemicals.

There are also computerised expert systems for predicting metabolism, either specifically or as part of general toxicity prediction (19–23). Their status with regard to chemicals testing has been discussed in the ECVAM report (3). Such programs can also be used to determine the log P values of metabolites and parent compounds, and to predict important properties of compounds, such as skin permeability. However, although they show great potential, they all require further improvement before they can be considered acceptable for routine application.

It is also suggested that maximum effort should be directed toward the further development and improvement of computer-based approaches to predicting metabolism.

## The potential of (Q)SAR and expert systems

The comprehensive ECVAM survey on alternatives (3) concluded that there is an urgent need for the validation and further refinement of the many existing (Q)SAR models and related computerised expert systems (24, 25). This is because the separate and combined application of these approaches offers the most realistic opportunity in the short-term and medium-term for providing useful information, which could contribute to substantially minimising the numbers of animals required, while allowing essential testing to proceed at a rate consistent with the requirements of the EU Chemicals Policy.

There are three main reasons why (Q)SARs and expert systems have not been used to their full potential: a) none have yet been formally validated; b) they need to be improved to cover a wider spectrum of toxic mechanisms of action, especially for endocrine disruption and non-genotoxic carcinogenesis (that are both based on receptor-binding); and c) their coordinated and combined use has not been explored sufficiently. Therefore, FRAME proposes, in conjunction with industry and academia, to address these issues in order to promote the widespread use of (Q)SAR and expert system approaches, particularly as a means of replacing and reducing the reliance on laboratory animals for the toxicity testing of new and existing chemical substances.

It should also be noted that, even when (Q)SAR and expert systems have not been formally validated, they can be used by industry for the in-house prioritisation of chemicals for testing, and to facilitate the development and evaluation of new alternative methods (26). Moreover, there is interest from regulatory bodies in using *in silico* approaches for chemical hazard identification, as illustrated in the work of Gerner and her colleagues at what is now the Federal Institute for Risk Assessment (BfR) in Berlin (27).

We stress that any initiatives aimed at further developing, improving and validating (Q)SAR and expert system approaches should take fully into account the ongoing activities of CEFIC, the EC Joint Research Centre (see "Projects and Activities" at <http://ihcp.jrc.cec.eu.int>), and the OECD. Moreover, the application of these approaches to risk assessment should be considered in relation to the potential for biokinetic modelling methods to generate information on the target organ concentrations of applied substances, under varying conditions of route of exposure and test species. There are several advances being made in this area, including the development of software and database programs for the rapid generation of new models. These will improve their utility for use in the process of testing large numbers of chemicals, and also for facilitating the interpretation of *in vitro* data for use in risk assessment (28; G. Loizou, personal communication). Previous suggestions for applying biokinetic modelling to hazard

identification have focused on drug development. However, recently Blaauboer has devised a scheme for its use in chemicals testing (29).

## Conclusions and Recommendations

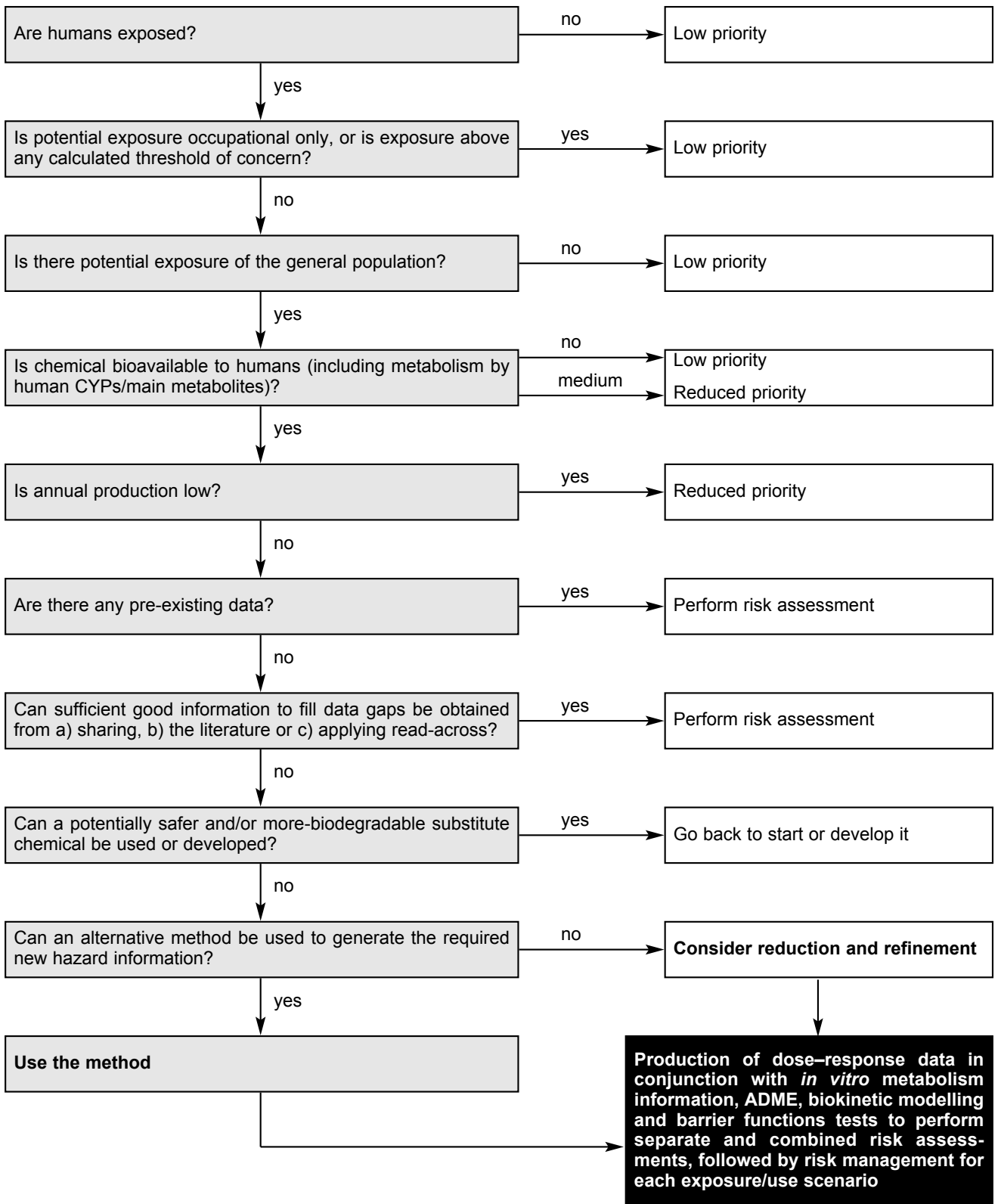
The scheme proposed in this paper (Figure 3) is intended to provide a framework for the efficient evaluation of general chemicals for their potential human hazard, while minimising the routine use of laboratory animal procedures. It is believed that this can be achieved, and indeed must be achieved, by the optimum use of non-animal approaches to testing, if the proposed timetable for the EU REACH system for new and existing chemicals is to be implemented sensibly and within a reasonable time.

The scheme is based on several fundamental premises, namely, that: a) toxicity testing should only be initiated once sufficient and relevant human exposure information has been obtained to permit a meaningful risk assessment; b) exposure should be predicted from a consideration of bioavailability, patterns of use and production levels; and c) any new testing should follow preliminary risk assessment, should involve the avoidance of duplicate testing and the maximum use of pre-existing data, should be driven by a justifiable need for specific kinds of data, and should be combined with a scientific and flexible approach to data acquisition. If maximum use is made of non-animal approaches for hazard identification and for prioritising chemicals, animal testing will only need to be undertaken as a last resort.

There is considerable scope for the increased use of non-animal methods for toxicity testing, not only involving methods that have received regulatory acceptance, but also those that are available for in-house testing. It is recommended that:

- The EU should embark on a period of public education about how risk assessment and risk management are conducted with respect to chemicals and chemical products.
- The proposed overall testing scheme should be used in conjunction with the ECVAM report on the status of alternative methods (3), especially with respect to the individual tier-testing schemes provided in that document.
- The timetable for developing and validating new test methods, presented in reference 3, should be actively implemented without delay, so that such methods are available in time for use in testing as many chemicals as possible. The following activities should also be priorities for future promotion by the EU: a) the improvement and validation of (Q)SAR models

**Figure 3: An overall decision-tree scheme for chemicals testing (see Table 1 for further details of each stage)**



and expert systems; b) the use of basal cell cytotoxicity as a predictor of acute lethal potency; c) the development and validation of a decision-tree system for general chemicals; d) the in-house use of skin and eye irritation studies (to distinguish between strong and weak irritants); e) the development of an integrated tier-testing scheme, involving computer predictions (of electrophilicity and metabolism), *in vitro* skin penetration, and release of IL-1 $\beta$  from dendritic cell cultures to predict contact sensitising potential; f) the *in vitro* prediction of non-genotoxic carcinogenicity via the use of cell transformation assays; and g) the development and validation of receptor-binding and cell culture methods for detecting endocrine disruptors.

- The general applicability of the threshold of regulation concept, especially to the testing of chemicals, should be further investigated.
- The improvement and validation of (Q)SAR models and expert systems is seen as absolutely crucial for the success of the new chemicals testing programme, since the potential of these methods is far greater than their actual applicability at the present time, which is limited by a lack of coordination of research effort, and also because they offer a realistic and economical way forward.

## Acknowledgements

A number of the ideas contained in this proposal arose in discussion during and following meetings of the ECVAM Working Group on Chemicals. We are grateful to the other members of the Group (Andrew Worth, Bas Blaauboer, Phil Botham, Johannes Doehmer, Julia Fentem, Manfred Liebsch and Horst Spielmann) for their collaboration, advice and comments on the manuscript.

## References

1. Anon (2001). *White Paper on a Strategy for a Future Chemicals Policy (COM(2001)88 final)*. Web site <http://europa.eu.int/comm/environment/chemicals/whitepaper.htm> (Accessed 8.1.01).
2. Combes, R.D. (2002). The EU white paper: *Strategy for a Future Chemicals Policy* (an idea that must have seemed good at the time!). *ATLA* **30**, 1–3.
3. Worth, A.P. & Balls, M., eds (2002). Alternative (non-animal) methods for chemicals testing: current status and future prospects. A report prepared by ECVAM and the ECVAM Working Group on Chemicals. *ATLA* **30**, Suppl. 1, 1–125.
4. Bottrill, K. (1998). The use of biomarkers as alternatives to current animal tests on food chemicals. *ATLA* **26**, 421–480.
5. Cramer, G.M., Ford, R.A. & Hall, R.L. (1978). Estimation of toxic hazard: a decision-tree approach. *Food and Chemical Toxicology* **16**, 255–276.
6. Atterwill, C.K. & Wing, M.G. (2002). *In vitro* pre-clinical lead optimisation technologies (PLOTs) in pharmaceutical development. *Toxicology Letters* **127**, 143–151.
7. Frawley, J.P. (1967). Scientific evidence and common sense as a basis for food-packaging regulations. *Food and Chemical Toxicology* **5**, 293–308.
8. Munro, I.C., Ford, R.A., Kennepohl, E. & Spreger, J.G. (1996). Correlation of structural class with no-observed-effect levels: a proposal for establishing a threshold of concern. *Food and Chemical Toxicology* **34**, 829–867.
9. Cheeseman, M.A., Machuga, E.J. & Bailey, A.B. (1999). A tiered approach to Threshold of Regulation. *Food and Chemical Toxicology* **37**, 387–412.
10. Combes, R.D. (1997). Defining a “threshold of regulation” the ultimate alternative for safety assessment? *ATLA* **25**, 187–191.
11. Kroes, R. & Kozianowski, G. (2002). Threshold of toxicological concern (TTC) in food safety assessment. *Toxicology Letters* **127**, 43–46.
12. Hanway, R.H. & Evans, P.F. (2000). Read-across of toxicological data in the notification of new chemicals. *Toxicology Letters* **116**, Suppl. 1, 61.
13. Barratt, M.D. (2000). Principles of toxicity prediction from 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, The Netherlands: Elsevier.
14. Kayser, D. & Roll, R., eds (1995). *Strategy of Toxicological Testing to Assess Chemicals within the Framework of the Chemicals Act*. BgVV Schriften (published in German and English). 76 pp. Munich, Germany: MMV Medizin Verlag.
15. Combes, R.D. (1992). The *in vivo* relevance of *in vitro* genotoxicity assays incorporating enzyme activation systems. In *Progress in Drug Metabolism, Vol 13* (ed. G.G. Gibson), pp. 295–321. London, UK: Taylor and Francis.
16. 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 hepatocyte and metabolically competent systems task force report 1. *ATLA* **27**, 579–638.
17. 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 engineering cell lines: characterisation and applications in toxicity testing. Report and recommendations of ECVAM workshop 26. *ATLA* **25**, 625–639.
18. Coecke, S., Bogni, A., Langezaal, I., Worth, A., Hartung, T. & Monshouwer, M. (2001). The use of genetically engineered cells for assessing CYP2D6-related polymorphic effects. *Toxicology in Vitro* **15**, 2553–2556.
19. 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.
20. Greene, N., Judson, P.N., Langowski, J.J. & Marchant, C.A. (1999). Knowledge-based expert sys-

- tems for toxicity and metabolism predictions: DEREK, StAR and METEOR. *SAR & QSAR in Environmental Research* **10**, 299–314.
21. Combes, R.D. & Judson, P. (1995). The use of artificial intelligence systems for predicting toxicity. *Pesticide Science* **45**, 179–194.
  22. Langowski, J. & Long, A. (2002). Computer systems for the prediction of xenobiotic metabolism. *Advanced Drug Delivery Reviews* **54**, 407–415.
  23. Darvas, F., Szabó, I., Karancsi, T., Slégel, P., Dormán, G., László, Ü. & Krajcsi, P. (2002). Dual uses of *in silico* and *in vitro* metabolism data in lead discovery, ComGenex' MAID is a metabolism alerting system for the early phase of drug discovery. *Genetic and Engineering News* **22**, 32–34.
  24. Worth, A.P., Barratt, M.D. & Houston, J.B. (1998). The validation of computational prediction techniques. *ATLA* **26**, 241–247.
  25. Worth, A.P. & Cronin, M.T.D. (2003). Report of the workshop on the validation of QSARs and other computational prediction models. *Proceedings of the Fourth World Congress on Alternatives and Animal Use in the Life Sciences, New Orleans, 2002*. *ATLA* **31**, Suppl. 2, in press.
  26. Barratt, M.D., Castell, J.V., Chamberlain, M., Combes, R.D., Dearden, J.C., Fentem, J.H., Gerner, I., Giuliani, A., Gray, T.J.B., Livingstone, D.J., Provan, W.M., Rutten, F.A.J.J.L., Verhaar, H.J.M. & Zbinden, P. (1995). The integrated use of alternative approaches for predicting toxic hazard. The report and recommendations of ECVAM workshop 8. *ATLA* **23**, 410–429.
  27. Zinke, S., Gerner, I. & Schlede, E. (2002). Evaluation of a rule base for identifying contact allergens by using a regulatory database: comparison of data on chemicals notified in the European Union with “Structural Alerts” use in the DEREK Expert System. *ATLA* **30**, 285–298.
  28. 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.
  29. Blaauboer, B.J. (2002). The applicability of *in vitro*-derived data in hazard identification and characterisation of chemicals. *Environmental Toxicology and Pharmacology* **11**, 213–225.
  30. European Centre for the Ecotoxicology and Toxicology of Chemicals (2002). Outline concept of a pragmatic approach to risk assessment. Web site [www.ecetoc.org](http://www.ecetoc.org) (Accessed 3.8.02).
  31. The Oil Companies' European Organisation for Environment, Health and Safety (2001). Suggested approach to the basic ranking of exposure under the REACH scheme. Web site [www.concawe.be](http://www.concawe.be) (Accessed 20.11.02).