

13th Meeting of the Scientific Group on Methodologies for the Safety Evaluation of Chemicals (SGOMSEC): Alternative Testing Methodologies and Conceptual Issues

Bas J. Blaauboer,¹ Michael Balls,² Martin Barratt,³ Silvia Casati,² Sandra Coecke,² Mostafa K. Mohamed,⁴ Jack Moore,⁵ David Rall,⁶ Kirk R. Smith,⁷ Raymond Tennant,⁸ B.A. Schwetz,⁹ William S. Stokes,⁸ and Maged Younes¹⁰

¹RITOX, Utrecht University, Utrecht, The Netherlands; ²ECVAM, European Commission, Ispra, Italy; ³Unilever Environmental Safety Laboratory, Bedford, England; ⁴Department of Community Medicine, Ain Shams University, Cairo, Egypt; ⁵Institute for Evaluating Health Risks, Washington, DC; ⁶National Institute of Environmental Health Sciences, Washington, DC; ⁷School of Public Health, University of California at Berkeley, Berkeley, California; ⁸National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina; ⁹National Center for Toxicological Research, Jefferson, Arkansas; ¹⁰International Programme on Chemical Safety, World Health Organization, Geneva, Switzerland

Substantial world-wide resources are being committed to develop improved toxicological testing methods that will contribute to better protection of human health and the environment. The development of new methods is intrinsically driven by new knowledge emanating from fundamental research in toxicology, carcinogenesis, molecular biology, biochemistry, computer sciences, and a host of other disciplines. Critical evaluations and strong scientific consensus are essential to facilitate adoption of alternative methods for use in the safety assessment of drugs, chemicals, and other environmental factors. Recommendations to hasten the development of new alternative methods included increasing emphasis on the development of mechanism-based methods, increasing fundamental toxicological research, increasing training on the use of alternative methods, integrating accepted alternative methods into toxicity assessment, internationally harmonizing chemical toxicity classification schemes, and increasing international cooperation to develop, validate, and gain acceptance of alternative methods. — *Environ Health Perspect* 106(Suppl 2):413–418 (1998). <http://ehpnet1.niehs.nih.gov/docs/1998/Suppl-2/413-418blaauboer/abstract.html>

Key words: risk assessment, alternative methods, toxicology, *in vitro* methods, reduction, refinement, replacement, testing, QSAR, PBBK modeling, transgenic models

Introduction

Countries commit considerable resources to determine the potential of chemicals to cause adverse effects to humans or their environment. Such efforts apply to chemicals

used as drugs, pesticides, food additives, cosmetics, and industrial processes, or those found in consumer products. Specific approaches incorporate a series of tests,

commonly involving laboratory animals, to first ascertain the potential of a chemical to cause an array of toxic effects, e.g., ocular toxicity, infertility, nervous disorders, cancer, etc. The array of toxic end points, and relevant protocols to detect such effects, are usually determined by regulatory agencies, based on the recommendation of their own scientists or ad hoc groups of scientific experts. Thus, the nature of testing evolves with advances in scientific knowledge; new end points may be identified or refinements proposed for methods used to assess previously selected end points. The overall goal is to devise strategies that prevent adverse impacts on the health of the public and the environment.

Alternative Test Methods for the Protection of Human Health and the Environment

Risk assessment refers to a structured sequence of analyses by which one reviews and characterizes the potential toxic properties of a chemical. Such analyses usually incorporate separate but integrated judgments based on qualitative and quantitative criteria. The qualitative process identifies the type and quantity of information needed and then reviews and integrates that information to reach a judgement as to the potential of a chemical to cause adverse effects. In rendering a qualitative judgment, one determines that there are adequate data to conclude that an adverse effect has been observed or could be anticipated and, if not in the species of concern, that the effect may occur in the subjects of concern, i.e., humans, ecosystems, etc. The subsequent step employs procedures that express the potential of the chemical to cause such effects at specified levels of exposure, with particular focus on the lowest doses that cause such effects and the changes in incidence or severity of the effects with increasing dose.

Many of the toxicology tests in common use today were designed and modified according to clinical and physiological criteria selected on the basis of the commonality of the observation, e.g., organ dysfunction, cancer, infertility, or birth defects, rather than understanding the underlying pathobiology of the process. Thus, tests were empirical and acceptance was based on correlation with observations in humans.

Alternatives to these empirical tests have been developed to recapitulate critical steps or events in the disease process leading to

This paper was prepared as a joint report for the 13th Meeting of the Scientific Group on Methodologies for the Safety Evaluation of Chemicals (SGOMSEC): Alternative Testing Methodologies held 26–31 January 1997 in Ispra, Italy. Manuscript received at *EHP* 9 May 1997; accepted 20 February 1998.

Address correspondence to Dr. B.J. Blaauboer, RITOX, Utrecht University, PO Box 80.176, 3508 TD Utrecht, The Netherlands. Telephone: 31 30 2535400 Fax: 31 30 253 50 77. E-mail: b.blaauboer@ritox.dgk.ruu.nl

Abbreviations used: DEREK, deductive estimation of risk from existing knowledge; ECVAM, European Centre for the Validation of Alternative Methods; MAPREC, molecular analysis by PCR restriction enzyme cleavage; NIEHS, National Institute of Environmental Health Sciences; PBBK, physiologically based biokinetic; PCR, polymerase chain reaction; QSAR, quantitative structure–activity relationships; RITOX, Research Institute of Technology; SAR, structure–activity relationships; SGOMSEC, Scientific Group on Methodologies for the Safety Evaluation of Chemicals.

the adverse response. These events may represent the mechanism of action or one of a cascade of events, often referred to as the mode of action, between exposure and the expression of disease. The fidelity of these isolated events for prediction of the end-stage disease ranges from poor to excellent. A desired property of alternative tests, or batteries thereof, is that they should support decisions as to toxicity on the basis of the composite of information. Individual test results should contribute to a profile of data such that the confidence in a test result can be enhanced or minimized on the basis of corroborative findings. Thus, test results can be used in a correlative sense to recapitulate the expected results from a test system such as a whole animal with all physiological processes intact.

Advances in general scientific knowledge may indicate the need to either expand the breadth of data collected or refine the type of data sought. Refinements in test procedures result in the collection of data that provides a better understanding of the factors associated with the development of an effect, rather than merely recording the presence and severity of such effect.

Over the years there has been a growing ethical and legal commitment to modify test strategies so as to result in the use of fewer animals as well as to modify procedures to be more sensitive to the welfare of the animals used in testing. Tiered approaches have been developed that are more efficient in cost and time and that also result in reduction, refinement, or replacement of animals in test schemes. Although visible progress has occurred in regulatory agencies, advances of at least similar magnitude have also been realized in commercial settings. Alternative methods have been used to predict the toxicity of synthesized intermediates or degradation products during the development of new products. Such applications benefit worker safety and the public by highlighting the need to limit exposure or by deciding not to develop certain material for commercial use. This represents disease prevention—the preferred public health goal.

In developing alternative test methodologies, one must be responsive to changes in the type of test data that are of value to the risk assessment process and to the data that provide the best guidance for product development. A strategy might entail the development of a series of *in vitro* tests that model the steps leading to the emergence of toxic effects in humans. Simple steps will be explored at first, leading to the development

of systems of increasing sophistication. The success of simple systems, e.g., receptor-binding assays, suggests that a series of tests could be successfully developed, perhaps in a tiered manner. Such an approach may have particular utility in refining or supplanting repeat-dose toxicity tests or even the complexities of carcinogenicity or developmental toxicity testing.

The development of a successful alternative test must go beyond the definition of a test protocol that provides consistency of results of appropriate specificity and sensitivity. It must also include a process in which test data can be converted to a form that can be used in the assessment or characterization of the toxic end point of interest. This may take a variety of forms ranging from a classification scheme to a measure of dose response, e.g., potency. Bruner et al. (1) have proposed the use of statistical methods to evaluate the predictive potential of assay methods such as ocular irritation or skin sensitization. Bristol et al. (2) have used an empirical approach to the prospective evaluation of methods proposed to predict potential carcinogens.

Progress in Development and Acceptance of Alternative Methods

Advances in new technologies and understanding of biological and chemical phenomena have enhanced the development and acceptance of alternative test methods over the past few years. The development of new tests continues to be driven by considerations of scientific credibility, the concept of reduction, refinement and replacement of animal use, and the desire for faster and less expensive methods. New test methods often reflect new knowledge of molecular biology and an increased level of scientific usefulness. The process for development of a new test method is more scientifically rigorous than in the past, resulting in increased likelihood that new tests will improve the quality of risk assessment.

Data from new tests are more likely to be useful to help build quantitative structure-activity relationships (QSARs) or other models because of better understanding of the scientific basis for the tests. Importantly, QSAR models thereby become more predictive of biological activity and can be the basis for developing hypotheses about mechanisms of toxicity. Because of better understanding of the biological basis of new tests, they are more easily integrated into predictive strategies and risk assessment strategies. Development of

new test methods has extended into areas of toxicology beyond those of earlier focus, that is, skin and eye irritation and skin corrosion. Alternative tests for prediction of carcinogenicity have received particular attention in recent years. Thus, new tests are developed today on the basis of improved mechanistic understanding in toxicology, the need for quantitative data for hazard identification and risk assessment, the need for test results to contribute to a profile of information for risk assessment, and the need for data to be extrapolated to the species of choice.

The increased use of computer software programs in areas of computational chemistry and various statistical and modeling procedures has made it possible to relate many aspects of chemical structure to observed biological and toxicological effects. It also made these tools more widely available. This has impacted the development of QSAR and other models that relate chemical exposure levels to tissue concentrations of toxicants and the associated toxic effects.

Mechanistic Approaches

There is currently much talk of the need for mechanistic tests, although 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.

A mechanism has been defined as an explanation of an observed phenomenon that explains the processes underlying the phenomenon in terms of events at lower levels of organization (3). Thus, a mechanistic test is based on a system at an acceptable level of organization and a relevant end point based on a sufficient understanding of the cellular and/or molecular basis of the effect under consideration. An example is a test based on interaction with a defined receptor, which is a critical or pivotal stage in the development of an effect.

Fidelity, Discrimination, Analogy, Mechanism, and Correlation. Fidelity is the accuracy with which a model reproduces the overall properties of what is being modeled, whereas discrimination is the accuracy with which a model reproduces a particular property or properties of what is being modeled. No model can offer 100% fidelity or 100% discrimination, but the best models will have the highest possible fidelity in combination with the highest possible discrimination. In general, a low fidelity/high discrimination model is more

likely to be useful than a high fidelity/low discrimination model (4).

The assumed relevance of animal tests to humans is based on the general high fidelity of animal models, i.e., on analogy (where similarity in a particular circumstance is inferred from agreement or similarity in an acceptable number of other features in the systems being compared), and not on mechanism (where similarity is based on an adequate knowledge of the mechanistic basis of the phenomenon under consideration and its operation in the systems being compared). In any case, similarity does not mean identity, so judgment in the interpretation of the meaning of data will always be necessary, whatever the model may be.

Correlative approaches, based solely on statistical relationships between phenomena that cannot be explained on a mechanistic basis, are unlikely to lead to correlative nonanimal tests that will receive widespread acceptance. This will apply even where such tests would be more useful than existing animal tests that also lacked a sufficient mechanistic basis.

Some qualitative and quantitative structure-activity relationship (SAR) models generally represent correlative approaches; but mechanistic SAR approaches are also being developed, e.g., when interactions with specific receptors can be predicted from structure and the consequences of such interactions are understood (see also the section on QSAR).

High Fidelity and Mechanistic Tests.

There can be confusion over whether a high fidelity test is a mechanistic test. For example, the use of whole rat embryos *in vitro* is an example of a high fidelity model, since the cultured embryos are very similar to rat embryos *in utero*. Thus, when whole embryo cultures are used to screen chemicals for teratogenicity (according to a number of specified, relevant end points), we have a high fidelity test, but we do not have a sufficient understanding of the cellular or molecular basis of teratogenicity for this to be termed a mechanistic test.

Mechanistic tests are the tests that are most likely to be high discrimination tests, but the fidelity of the system must also be borne in mind. For example, the *Salmonella typhimurium* test is considered to be a relatively high discrimination test for genotoxicity, but a liver S9 fraction must be incorporated to improve its fidelity, i.e., its ability to detect metabolism-mediated genotoxicity.

It is commonly believed that validation is the limiting step in the acceptance of

new test methods. However, it is now becoming clear that, in fact, the main limiting factor is new test development. The development of relevant and reliable mechanistic toxicity tests must depend on the rate of progress achieved in the fundamental science of toxicology.

Existing Tests, New Tests, and Knowledge Needed. Because the data from an animal test are themselves of limited usefulness in terms of the purpose of the test, e.g., for predicting particular likely effects in human beings, those data must also be of limited utility as a basis for evaluating the reliability of *in vitro* test data for predicting the likelihood of those effects in human beings.

It is essential to have a regular, thorough, and objective review of all test methods in light of the purposes for which they are used. This, in turn, requires an objective analysis of those purposes. If it is our aim to use tests to provide essential knowledge as a means of developing the safest, most effective products, we must first define more precisely the knowledge needed to make this possible. If another aim is to develop valid nonanimal test procedures, we must decide how these new test procedures should be validated. If the existing animal test can be shown to be reliable and relevant in providing the knowledge needed, then data from that test can be used in the validation of potential replacement alternative methods. If not, then no attempt should be made to use such data in the validation of new tests. In those circumstances, the way forward is to establish a convincing relationship between the information that can be provided by the nonanimal test procedure and the knowledge needed to predict likely effects in human beings.

Another problem in new test development and validation is related to the availability of sufficient high quality data about the *in vivo* effects of an adequate number and range of chemicals. Some biological data, particularly data generated prior to the advent of Good Laboratory Practices (GLP), do not meet today's stringent requirements for acceptability. In some cases it may be possible to use QSARs to upgrade these data so that they are acceptable for use in the development and validation of alternative tests. If QSAR techniques can be used to demonstrate that the results of these tests are consistent with the physicochemical attributes of the chemicals when compared with the results from tests conforming to the current acceptance criteria, they should be acceptable for

use in the development and validation of *in vitro* alternative methods (4).

Use of Chemical Parameters and Computer Algorithms

The development of computer-based methods of assimilating and analyzing diverse chemical properties provides the opportunity to create algorithms for predicting toxicological effects of congeneric chemicals. One method that has been extensively developed is QSAR based on the premise that the properties of a chemical are implicit in its molecular structure. As a consequence, if a mechanistic hypothesis can be proposed linking a group of related chemicals with a particular toxic end point, the hypothesis is then used to define relevant parameters to establish a structure-activity relationship. The resulting model is then tested, and the hypothesis and parameters refined, until an adequate model is obtained. For a QSAR to be valid and reliable, the dependent property for all of the chemicals covered by the relationship has to be elicited by a mechanism that is both common and relevant to that dependent property (5).

The principles of QSARs also need to be applied to the development of *in vitro* alternatives to animal tests if those methods are to be reliable. Historically these principles have been overlooked in many cases with unfortunate results. Some alternative tests determine end points that are substantially different from those they claim to predict because the mechanism modeled by the *in vitro* alternative represents only part of that which is active *in vivo*. In other cases, tests have been developed that can predict end points accurately for some classes of chemicals, but are then wrongly assumed to be applicable to all chemical classes. The fact that different types of chemicals may elicit changes in a particular biological end point via different mechanisms has clearly not been appreciated.

It has long been recognized that for a chemical to be biologically active, it must first be transported from its site of administration to its site of action (partition) and it must then bind to or react with its receptor or target (reactivity). If any QSAR or *in vitro* model is deficient in modeling either partition or reactivity, only a partial correlation with the *in vivo* response is likely to be observed, e.g., the varying degrees of partial correlation with *in vivo* data found with the many *in vitro* methods that have been developed and advocated as alternatives to the Draize rabbit eye irritation test (6). Thus, it follows

that for an *in vitro* test to reliably predict *in vivo* toxic potential, it should be sensitive to the same parameters that are responsible for the effects *in vivo*; such a test would be expected to show a high degree of correlation with the response *in vivo*.

One example of the use of a mechanistic approach in SAR is provided by the expert system DEREK (the deductive estimation of risk from existing knowledge). DEREK uses rules based on correlations between the structure of chemicals and their toxicological activity, supported by the knowledge of organochemical reaction mechanisms. Examples of physicochemical-based QSARs are the models for skin corrosivity of organic acids, bases, phenols, and electrophiles (5).

A recurring feature of QSAR models for the classification of toxicological hazard is the problem of biological uncertainty at boundary regions. The concept of the boundary region has its origin in the fact that most regulatory schemes operate initially by quantizing continuous biological (toxicological) data into discrete hazard bands that can conveniently be used in the regulatory process. It is the biological variability inherent in toxicological testing that leads to uncertainty in classification in the boundary regions. This variability could manifest itself as the results of two well-conducted Draize rabbit eye irritation tests on the same chemical leading to a nonirritant classification in one case and an irritant classification in the other. Away from the boundary region, the inherent biological variability is less likely to result in two separate tests leading to different classifications. QSAR techniques such as principal components analysis afford visualization and hence predictability of regions of chemical parameter space in which ambiguity in *in vivo* results may arise.

There are three ways in which to select a set of chemicals for the validation of an alternative method:

- Make a selection of chemicals that guarantees success (all of the chemicals are far away from classification boundaries).
- Make a selection that guarantees failure (all of the chemicals are on or close to a classification boundary).
- Select the chemical objectively, by trying to retain balance among the various mechanistic types, between classification categories, and between classification categories in each mechanistic type.

The use of principal components mapping allows the selection of chemicals that cover the widest possible parameter space in terms of both biological activity

and physicochemical properties. Techniques of this type have been used in connection with the selection of test chemicals for the ECVAM-sponsored study on skin corrosivity (7).

Estimating Exposure-Effect Relationship (Physiologically Based Biokinetic Models)

Understanding the absorption, distribution, metabolism, and excretion (i.e., the biokinetics) of toxic chemicals is necessary to predict the target organ concentrations of the chemical or its active metabolite(s). Putting this dynamic information into a mathematical framework, or model, allows the prediction of target organ concentrations of chemicals after exposure. In physiologically based kinetic modeling a number of physiological parameters (such as blood flows, organ volumes) and chemical-specific parameters (e.g., tissue solubility, biotransformation rates) are combined, which allows the prediction of organ or tissue concentrations of chemicals, given a certain external dose. These model parameters can in many cases be derived from *in vitro* data or from calculations on the basis of physicochemical properties of the compound under study and its metabolite(s). This is especially the case for the calculation of tissue-blood partition coefficients. A number of examples can be given for the calculation of critical tissue concentrations on the basis of physiologically based biokinetic models (PBBK) (8).

PBBK allows the integration of data derived from relevant *in vitro* toxicity tests in an assessment of a compound's systemic toxicity. If, for instance, the neurotoxicity of a compound can be measured in cell cultures, the minimal effective concentration for the compound can be used as the input for the tissue concentration in the PBBK model. This will then allow the calculation of the corresponding dose. However, such an approach implies that the relevant mechanism of toxicity is present in the *in vitro* system.

Other important advantages of PBBK models are the use of human blood levels of chemicals in an exposure-risk assessment and their application to interspecies extrapolation. Because a model can be constructed for any species for which the physiological parameter (blood flow, organ volume, etc.) is available, data empirically determined in one species can be translated to another. This offers the possibility to make a more realistic choice of the species studied, leading to a better design

of toxicological studies. Similarly, extrapolations for different dose ranges and routes of administration of chemicals are possible. Even if not all necessary data are available based on *in vitro* or other nonanimal-based data, the application of models partly based on experiments with a limited number of animals is of great value in reducing the number of animal studies needed. It also can be helpful in the most economical design of animal studies by giving the appropriate concentrations to be studied (8).

Application of Molecular Biology

New methods in the field of molecular biology have permitted the development of animal models with genetic alterations that are specific for new test methods. This involves introducing or deleting genetic material to make the animal resemble the human more closely or to make the animal more sensitive and specific for a type of toxicity such as carcinogenicity (9). Significant progress has been achieved over the past two decades in defining elements of chemical structure and results of genotoxicity tests that predict chemical-induced carcinogenicity. Although much has been achieved, both product development and public health decisions rely upon obtaining bioassay results. The search must continue for more effective methods of assessing the carcinogenic potential of substances, which can contribute to safety assessment decisions. The rapid and substantial rate of progress in molecular and cell biology and genetics has resulted in both new knowledge of specific genes that are involved in the induction and development of cancers in rodents and humans, and of techniques using such genes as specific targets for the action of potential carcinogens. These advances in knowledge will continue to provide new models and systems to improve drug and chemical safety assessments.

Recent developments in methodologies to complement or supplant long-term carcinogenicity bioassays include transgenic mouse models. These models use unique phenotypic properties imposed either by pronuclear injection of specifically regulated oncogenes or via knock-out of a tumor-suppressor gene. Properties of the transgenic models that make them most appropriate as bioassays are as follows:

- They provide for appropriate disposition and metabolism of test agents: they provide specific genotypic targets and phenotypic responses to exposures.

- They provide the end point of direct interest (e.g., tumor development).
- Single species can minimize the consequences of strain- or species-specific effects encountered in long-term conventional bioassays by significantly reducing the time to response to 6 months.
- The influence of spontaneous tumors is minimized.
- The short duration of exposure and uniformity of responses allow for a significant reduction in the number of animals needed to perform a bioassay (9).

Because significant advances in molecular biology and transgenic models continue at a rapid pace, the current models may be viewed as prototypes, but their use can provide valuable information on the properties of specific chemicals as well as providing a database with which any future transgenic models can be efficiently compared. The *p53*-deficient, Tg.AC, and *ras*HII models are currently being evaluated through a multilaboratory international effort that will result in a database of 40 to 50 chemicals by the end of 1997. The evaluation effort involves the use of known positive controls tested in all participating laboratories, together with unique agents tested in selected laboratories, but with a group of carcinogenic and noncarcinogenic substances tested in common in all three models. Since the chemicals are drawn mostly from substances that have undergone rodent carcinogenicity bioassay by the U.S. National Toxicology Program, a direct comparison within and between the assays is achieved. In addition, chemicals undergoing conventional carcinogenicity bioassays can be used to prospectively evaluate the transgenic models. Because a mechanistic basis can be inferred for the transgenic models, information on chemical structure, genetic, or systemic toxicity, or other properties can be used to predict the outcome of the prospective assay. It is also possible to analyze tumors induced in the transgenic lines for transgene expression or for the induction of mutations. Such data can contribute to both further verification of the models as well as provide additional information on the specific action of the chemical or drug.

Implementation Considerations

Complexity of Approaches and Availability of Methods

The trend toward alternative methods will undoubtedly lead to a wider variety of

approaches to hazard prediction and risk assessment. This will be due to the much greater range of technologies and methods that will have to be applied. Often, data provided by various methods will have to be combined and integrated as contributions to the overall decisionmaking process. This, in turn, will lead to the need for a wider range of equipment, expertise, and experience. These new approaches will be effective only if they result in integrated, stepwise, and tier-testing strategies, aimed at giving the most relevant, reliable, and useful outcomes quickly and inexpensively (10).

These new methods, new skills, and new strategies will require training programs to provide personnel, and data banks to provide ready access to test protocols, prediction models, outcomes of validation studies, and evidence of experience in use. Ideally, this approach would be based on international cooperation, leading to the greater harmonization of test guidelines and to agreement on the principles of validation and acceptance of alternative methods.

The development of unnecessary tests must be avoided. The need for a new test in relation to other tests that could provide the same sort of information should be satisfactorily established at an early stage of test development.

Integration into a Strategy

In a toxicological risk assessment it is logical to take into consideration information from all available sources. How this is done is highly dependent on the goal of the risk assessment process. In many cases it is subject to the judgment of the person conducting the exercise. Such information can include animal data, epidemiological data, mechanistically based information derived from *in vitro* models, or computer-generated data. Examples from several of these approaches are cited in accompanying documents (4,5,8,9).

For the introduction of new methodologies, it is important to clarify how the methods can be incorporated into an integrated approach. For example, a test employing cell cultures will not be easily interpreted in terms of the systemic toxicity of a compound without taking into consideration the biokinetics of the compound or its metabolites. It will thus be possible to link the external dose with the appropriate *in vitro* toxicity data and make predictions on the toxicity of the compound under study. Taken together, these approaches will result in test strategies that lead to prediction of a compound's toxicity

while reducing the reliance on strategies solely based on the use of animals.

Criteria of Acceptance: Lessons from Examples

An example of the application of advances in molecular biology to alternative methods is the development of two methods for neurovirulence testing of modified live oral polio vaccines. For more than 40 years, monkeys have been used to test batches of this vaccine for neurovirulence. This test is expensive, labor intensive, and requires a large number of animals. The World Health Organization (WHO) has completed a collaborative study on the use of the molecular analysis by polymerase chain reaction (PCR) and restriction enzyme cleavage (MAPREC) assay. This assay is based on quantification of revertants at position 472 of the 5' noncoding region of the poliovirus genome, which has been shown to produce neurovirulence in the monkey test (11). Studies have demonstrated the usefulness of the MAPREC assay as a screening test to predict neurovirulence. Positive results in this assay are now considered to be predictive of neurovirulence and therefore eliminate the need for additional testing in monkeys.

A transgenic mouse model (Tg PVR) that is susceptible to the poliovirus has been developed by introduction of the human gene that codes for the cellular receptors to poliovirus into the mouse genome. This mouse model develops clinical signs and morphological lesions in the central nervous system similar to those in primates when infected with neurovirulent poliovirus strains. Recent studies (11) have demonstrated that this mouse model is as sensitive as the monkey test and can be considered a potential replacement for the monkey test. This example further demonstrates the potential for applying molecular biology techniques to new alternative test methodologies.

Summary

Substantial resources are being committed worldwide in the search for alternatives to the use of animals for the protection of human health and the environment. Such commitment demands strong scientific stewardship of resources. Concomitantly, it must be recognized that the search is an ongoing process that must intrinsically be driven by new knowledge emanating from fundamental research in toxicology, carcinogenesis, molecular biology, biochemistry, computer sciences, and a host of other

disciplines. Progress is best achieved through international cooperation and harmonization that is based upon critical evaluation and strong scientific consensus. In this way the most useful alternative methods will emerge for the safety assessment of drugs, chemicals, and environmental factors.

Recommendations

- The rate of development of alternative tests for use in toxicological assessment should be increased.
- An increased emphasis should be placed on the development of mechanism-based methods for specific aspects of toxicity.
- Investment should be increased in the development of fundamental research that underpins toxicology and toxicity testing.
- Training institutions, granting bodies, and regulatory agencies should be encouraged to support research and training that will enhance the development and use of alternative systems.
- Accepted alternative methods should be integrated into toxicity assessment of chemicals.
- International cooperation in development, validation, and acceptance of alternative methods should be encouraged.
- In the interest of the most effective development and use of alternative methods, international harmonization of chemical toxicity classification schemes should be encouraged.

REFERENCES AND NOTES

1. Bruner LH, Carr GJ, Chamberlain M, Curren RD. Validation of alternative methods for toxicity testing. *Toxicol in Vitro* 10:479-501 (1996).
2. Bristol DW, Wachsmann JT, Greenwall A. The NIEHS predictive-toxicology evaluation project. *Environ Health Perspect* 104(Suppl 5):1001-1010 (1996).
3. Frazier JM. The role of mechanistic toxicology in test method validation. *Toxicol in Vitro* 8:787-791 (1994).
4. Balls M. Commentary: Mechanistic approaches and the development of alternative toxicity test methods. *Environ Health Perspect* 106(Suppl 2):453-457 (1998).
5. Barratt MD. Integration of QSAR and *in vitro* toxicology. *Environ Health Perspect* 106(Suppl 2):459-465 (1998).
6. Bagley DM, Brunner LH, de Silva O, O'Brien KAF, Uttley M, Walker AP. An evaluation of five potential alternatives *in vitro* to the rabbit eye irritation test *in vivo*. *Toxicol in Vitro* 6:275-284 (1992).
7. Barratt MD. Quantitative structure activity relationships for skin corrosivity of organic acids, bases and phenols. *Toxicol Lett* 75:169-176 (1995).
8. Blaauboer BJ, Bayliss, MK, Castell, JV, Evelo, CTA, Frazier, JM, Groen, K, Gulden, M, Guillouzo, A, Hissink, AM, Houston, JB, et al. The use of biokinetics and *in vitro* methods in toxicological risk evaluation. *ATLA* 24:473-497 (1996).
9. Tennant RW. Evaluation and validation issues in the development of transgenic mouse carcinogenicity bioassays. *Environ Health Perspect* 106(Suppl 2):473-476 (1998).
10. Blaauboer BJ, Balls M, Bianchi V, Bolcsfoldi G, Guillouzo A, Moore GA, Odland L, Reinhardt CA, Spielmann H, Walum E. The ECITTS integrated toxicity testing scheme: the application of *in vitro* test systems to the hazard assessment of chemicals. *Toxicol in Vitro* 8:845-846 (1994).
11. Dragundsky E, Tafts R, Chernokhvostova Y, Nomura T, Hioki K, Gardner D, Norwood L, Levenbook I. A Poliovirus-susceptible transgenic mouse model as a possible replacement for the monkey neurovirulence test of oral poliovirus vaccine. *Biologicals* 24:77-86 (1996).