SUMMARY OF CONSIDERATIONS IN THE REPORT FROM THE OECD EXPERT GROUPS ON SHORT TERM AND LONG TERM TOXICOLOGY

Test Guidelines for Toxicity Testing were prepared in draft form by the two OECD Expert Groups on Short Term Toxicology and Long Term Toxicology (List of Participants appended). The drafts were issued in December 1979 from the Lead countries (U.K. and U.S.A.) in a final report, which included detailed considerations of the Groups' approaches, objectives and general principles in toxicity testing, etc. The following text constitutes a summary of the considerations and is presented as background information to users of OECD Test Guidelines.

GENERAL

The aim of the present work has been to produce a framework for each toxicity test which is sufficiently well-defined to enable it to be carried out in a similar manner in different countries and to produce results that will be fully acceptable to various regulatory bodies. The growing demands for testing and evaluating the toxicity of chemical substances will place increasing pressure on personnel and laboratory resources. A harmonized approach, promoting the scientific aspects of toxicity testing and ensuring a wide acceptability of test data for regulatory purposes, will avoid wasteful duplication or repetition and contribute to the efficiency use of laboratory facilities and skilled personnel.

When assessing the results of toxicological testing on any chemical, the limitations of the tests must be borne in mind. Of necessity, animals or in vitro systems are used and the results then extrapolated to man. Such extrapolations may not always be accurate. Futhermore, since it is not possible to devise "standard" test methods appropriate to all chemicals, judgement must be exercised in each case to assess the suitability of a particular method. In many areas of toxicology controversy exists concerning the appropriate experimental design to be used. For example, there are varying opinions on the duration of a study and the number and types of species of test animals which are considered necessary. Other limitations to the assessment of hazard to man derive from the fact that, as a rule, single substances are tested by defined routes

of exposure, whereas normal exposure is to a combination of substances, usually by more than one route of exposure. Also, it should be obvious that due to limitations of the size of animal experiments, low incidence effects may not be recognised. Consequently, although the results of toxicity testing will, in most cases, give good indications of possible hazard, they do not eliminate the need for continuing careful observations of humans.

The test guidelines will be adequate for the evaluation of most chemicals, and further elaboration or extension should only be done with good reason. Scientific judgement is essential in determining the conduct of a particular test so that a reasoned flexibility of approach is always necessary. The present guidelines have been developed taking into consideration a proper balance between resources and scientific requirements.

The Expert Groups stressed that more research is needed in order to improve experimental design and validation in these various fields of toxicology. These efforts may lead to tests which are more valid and also less demanding on resources in the assessment of the potential hazards posed to man by chemicals. As better methods are developed, it is essential that they replace, or complement, those recommended here.

A number of national and international documents were reviewed in the development of the final report. The review of these documents demonstrated both common elements as well as variations in approach and experimental design. As the development of the guidelines progressed, the use of draft guidelines and reports from a number of nations and international organisations were made available for consideration by the Groups. In the course of their review, the Groups found there were relatively few areas in toxicology where well-established and validated test methods could be easily incorporated in a harmonized approach to testing. As a result, a significant proportion of the Groups' activities, following the detailed review of the principles underlying toxicity testing, were devoted to the preparation of test guidelines which cover the key points of the test methods.

The test guidelines do not approach the level of detail found in standard operating procedures or similar documents. This is intentional because toxicology is a developing experimental science, and excessive rigidity or over-detailed specification of methods could inhibit scientific initiative and be counter-productive. There must be provision for the exercise of toxicological skill and judgement during the course of the study, even where this forms part

of a prescribed set of test requirements, and so guidelines or similar defined procedures should allow for this; obviously, the rationale for changes in procedure must be explained and supported scientifically. The emphasis on a flexible approach should not be construed as a recommendation for a lack of order. it should be seen as creating a situation in which the examination of the toxicity of a chemical substance is conducted as a scientific exercise rather than as a set of stereotyped tests to be conducted in a routine.

SIGNIFICANCE OF TOXICITY TEST AREAS REVIEWED

Acute Toxicity

This is examined to determine the degree of toxicity of a chemical substance, that is, the relationship between dose and adverse effects; to establish its toxicity relative to other chemical substances whose acute toxicity is known; to determine specific toxic effects; and to provide information on the mode of toxic action. A suitably designed acute toxicity study will also provide information from which a median lethal dose (LD50) can be calculated. By studying the effects, following administration by different routes, the relative hazards of different pathways of exposure can be assessed. By using animals of both sexes, sex differences in toxic response can be detected.

Acute toxicity studies will thus identify highly toxic chemicals and provide information on the possible hazards which could occur where humans are exposed. The slope of the dose response curve and the type of toxic response in experimental animals are of use in human health hazard evaluation; exposure to single acutely toxic doses of a chemical represents an abnormal or accidental situation for general human exposure. The numerical value of the median lethal dose (LD50) is widely used in toxicity classification systems, but it should not be regarded as an absolute number identifying the toxicity of a chemical substance. LD50 values for the same chemical may vary from study to study and between species or within a species because acute toxicity is influenced by both internal and external factors.

Short Term Repeated Dose and Subchronic Toxicity

While acute toxicity deals with the adverse effects of single doses, a more common form of human exposure to many chemical substances is in the form of repeated doses which do not produce immediate toxic effects. Delayed effects may occur due to accumulation of the chemical in tissues or other mechanisms, and it is important to identify any potential for these by subchronic testing.

The term "subchronic" has been used to embrace the toxic effects associated with repeated doses of a chemical over part of an average lifespan of experimental animals. The division between subchronic and chronic dosing regimes is sometimes taken as 10 per cent of the test animals' lifespan. Dosing periods lying between the single dose and 10 per cent of lifespan dosage are often called subacute. It was considered that this term was semantically incorrect and, therefore, to distinguish such dosing periods from the classical subchronic they may be described as "short-term repeated dose studies"; this applies to 14, 21 and 28-day studies. The main study durations involved have been 14, 28 and 90 days. Other study durations have been used in toxicology, but the selection of three primary durations, which have either the backing of experience or existing regulatory requirements, is considered to represent a reasonable approach. In general, the longer the subchronic study, the more information that is likely to be gained. This is an area in which a scientific comparison of the data from studies of different subchronic durations on the same chemicals should be carried out to determine their relative utility.

These studies will provide detailed information on toxic effects, target organs, reversibility or otherwise of effects and an indication of a "no effect level". The Group recommended that the use of a satellite group of test animals, given the highest dose and then observed after the ending of dosing, be considered to give additional information on the persistence or reversibility of effects.

These studies are important because they will be the first, and perhaps for some chemicals the only, repeated dose studies. It is therefore necessary to derive the maximum amount of information from them, and this is reflected in the extent of the guidance of clinical chemistry and histopathological investigations.

Local Effects on Skin and Eye

Determination of the surface effects of a chemical on the skin and eye is important because accidental contamination is always a possibility. Hazards may be related to physical form, with a liquid or particulate having a greater potential for contaminating the body surface and to physical properties, notably the pH, which can indicate a potential for producing extensive tissue damage or corrosion. If a chemical is found to be a powerful irritant or corrosive in skin studies, it is not considered necessary to repeat these studies in the eye as the effects will usually be even more marked. The only indication for eye studies in such cases is to determine the effects of treatment such as immediate eye washout.

Allergic Sensitisation

Allergic sensitisation, such as that which can occur following exposure by the dermal or inhalation routes, presents problems to significant numbers of humans, both in the occupational field and in the general population. Allergic reactions are of various types, but all involve at least one exposure to initiate the process of sensitisation. Early identification of any allergic potential is considered advisable to ensure that appropriate methods of control can be applied.

Reproductive Toxicity

The term covers the areas of reproduction, fertility and teratogenicity. It has been found that many chemicals can affect fertility and reproduction, often in an insidious manner without other overt signs of toxicity. Fertility can be affected in males and females, and effects can range from slightly decreased reproductive capability to complete sterility. Teratogenicity deals with the adverse effects of a chemical on the developing embryo and foetus. Reproductive toxicity is important as it has an important bearing on the health of mankind. Testing techniques are developing and the concept of combined tests, covering all aspects of reproductive toxicology, appears promising.

Carcinogenicity

The objective of a long-term carcinogenicity study is to observe test animals, for a major portion of their life span, for the development of neoplastic lesions, during or after exposure to various doses of a test substance by an appropriate route. Such an assay requires careful planning and documentation of the experimental design, a high standard of pathology and unbiased statistical analysis.

Chronic Toxicity

The objective of a chronic toxicity study is to determine the effects of a test substance in a mammalian species following prolonged and repeated exposure. Under the conditions of this test, effects which require a long latent period, or are cumulative, should become manifest. The application of these guidelines should generate data on which to identify the majority of chronic effects and to determine dose-response relationships. Ideally, the design and conduct should allow for the detection of general toxicity including neurological, physiological, biochemical effects and exposure-related, morphological effects.

Combined Chronic Toxicity/Carcinogenicity

The objective is to determine effects of a test substance which would be provided individually in carcinogenicity and chronic toxicity studies.

Toxicokinetics and Metabolism

Toxicokinetics is defined as the study of the rates of absorption, distribution, metabolism and excretion of toxic substances or substances under toxicological study. Metabolism is broadly defined as all aspects of the fate of a substance in an organism, and thus includes absorption, tissue distribution, biotransformations and excretion by all routes. The term toxicokinetics covers the rate of all the processes included under metabolism.

Data from toxicokinetic studies are desirable to aid in the evaluation of test results from other toxicology studies and in extrapolation of data from animals to man. Studies should be done on each chemical of toxicological concern. The concern may be predicated on the level and type of toxicity observed or anticipated and by the magnitude of potential human exposure.

Toxicokinetic studies also provide data useful for selecting appropriate dose levels for use in chronic toxicity and carcinogenicity studies by providing information about dose-dependent kinetics.

The time at which it is best to do a toxicokinetic study varies with the need for data to evaluate the safety of the test chemical. In certain cases, the initial experiments for determining absorption, distribution and excretion of the test chemical may be done soon after the acute toxicological studies. Further experiments, establishing the metabolic fate of the compound, may be needed for chemicals which will likely undergo chronic testing. If the results of toxicological studies indicate that further information on the metabolism of the test chemical is needed, identification and characterisation of major metabolites in blood and urine should be undertaken. For some purposes, dose-related toxicokinetic studies may be carried out. In pregnant animals, a kinetic analysis makes it possible to assess the amount of placental transfer of the parent compound, and its metabolites, at critical periods of organogenesis in relation to maternal exposure.

Step Sequence Testing

The Toxicology Groups included a report with recommendations for a Step Systems approach to hazard evaluation. The recommendations for a Minimum Premarket Data set (MPD) are expanded upon by providing indications for performing the tests and examples of exceptions from inclusion in MPD.

CONSIDERATIONS IN TOXICITY TESTING

Chemical Analysis

The identity of the test substance should be defined. The physical and chemical properties of the test substance provide important information for the selection of the route of administration, the design of the studies and the handling and storage of the test substance. It is important to characterise test mixtures and to identify impurities that are known or likely to be present. Separate studies of impurities may provide useful evidence in the evaluation of the carcinogenicity of the mixtures.

Choice of Test Animals

There is no experimental laboratory species which is identical to man in terms of structure or metabolism. There are obvious resemblances and similarities in function between man and other animal species, but even in the case of man's fellow primates, these are not such that straightforward extrapolations from animal tests to man are possible. The interpretation of animal test results in the assessment of possible human health hazard remains a matter of skilled judgement.

Accepting that no ideal animal analogue of man is available for laboratory testing, the choice of test species can be influenced by other considerations of a logistic nature, such as ease of breeding or purchasing, animal husbandry, speed of growth/development and handling under the experimental conditions. Rodents fulfil many of the logistic requirements and so are used extensively in toxicological studies. For acute oral, dermal and inhalation studies the rat is the preferred species with the option of the rabbit in the case of the dermal study. In the latter context the rabbit has the advantage of a larger size combined with a reasonable background of information on its behaviour in dermal studies, but, from the point of view of the comparison of toxic effects by different routes and the hazard evaluation, there is much to be said in favour of the rat, which will also have been the test species in studies by other routes.

In the eye and skin irritation and corrosivity testing, the guidelines are based on the methods developed by Draize, and the rabbit is thus the species of choice. For skin sensitisation the species used in the six recommended methods are guinea pig, rabbit, mouse and dog. In neurotoxicity studies, where the mechanism of action is by cholinesterase inhibition, the hen appears to be the most suitable test species.

It was generally agreed that both rodent and non-rodent species be included in the subchronic and chronic guidelines, recognising that a number of factors might dictate the number or choice of species for study.

In the subchronic studies, similar considerations are involved. There are cases where the use of a non-rodent species (in addition to a rodent species) is indicated to examine in greater depth toxic responses in different species. A species commonly used is the dog, and this has been covered in separate guidelines for oral studies. In subchronic dermal studies, the guinea pig is added to the recommended species on the grounds of logistic factors and an adequate background knowledge.

Historically, it has generally been recommended that chronic testing be perfomed with two mammalian species, one a rodent and another a non-rodent. The rat has normally been the rodent of choice. Of the non-rodents, dogs and primates have been most extensively utilised, due to their large size and their general availability. Moreover, the use of dogs and primates facilitates the performance of clinical and biochemical examinations. It should be noted that availability of non-rodents for research purposes may be a problem internationally. It should also be recognised that the lack of results with a non-rodent may impose a serious reduction in the sensitivity of the test to assess important effects which might be encountered in humans.

Such a dichotomy cannot be resolved in a generic manner. While it is still to be acknowledged that chronic effects obtained from both a rodent and non-rodent are needed, the selection of appropriate specie(s) for the chronic test may best be based upon practical reasons as well as the results of previously conducted tests. Testing with a single species may provide sufficient data for assessing the hazard of the chemical.

For carcinogenicity studies, a compound of unknown activity should be tested in both sexes in each of two animal species. Of the three rodent species of choice, the mouse and the rat have been more widely used than the hamster. The Syrian golden hamster has proved to be

useful in revealing carcinogenic effects, primarily in the respiratory tract and urinary tract; however, there is evidence that this species may be more widely used in the future for general carcinogenicity screening. Other species may be useful for special purposes.

The bioassay design should assure that variability of tumour incidence due to chance does not interfere in the interpretation of results. The variation of "background" tumour incidence should be adequately defined in the animal colonies used for carcinogenicity bioassays.

In all cases the animals should be healthy, of known origin, reasonably consistent in terms of age and body weight, and suitably acclimatised to the experimental environment before the study commences. In general, the guidelines do not specify specific age or weight ranges for test animals, but instead there are references, for example, to the use of young adult animals. Similarly, there is no firm recommendation for the use of specific strains as it is considered that at the present time the state of development of testing provides no firm justification for such a recommendation.

Animal Care

Stringent control of environmental conditions and proper animal care techniques are mandatory for meaningful results. Diet should meet all the nutritional requirements of the species used in the tests. It is highly desirable to know the effect of the dietary regimen on metabolism and on animal longevity as well as the development of toxicity. Variations in the use patterns of industrial and agricultural chemicals throughout the world preclude harmonization on one list of dietary contaminants. Notwithstanding this fact, common dietary constituents which are known to influence toxicity should not be present in interfering concentrations.

Number and Size of Groups

With the objective of an efficient approach to testing chemicals, there is no point in having more groups or more animals per group than are strictly necessary to attain the end-point of the reliable detection of toxic effects.

Toxicity studies are undoubtedly expensive in financial and resource terms. Part of the cost is related to the number of animals and the extent of clinical, necropsy and histopathological investigations required. Taking account of the inherent variability of biological systems, there must always be a balance between the number of animals theoretically required

to detect all effects from the weakest upwards and the number required to detect significant toxic effects. In a well-conducted study which goes according to plan it is possible to use small numbers of animals. However, in tests of chemicals of unknown toxic characteristics, problems often arise because the actual responses of animals differ widely from those anticipated when the study was designed. To deal with this problem it is prudent to increase the number of animals in order to ensure that sufficient animals are available at key points of a study to provide adequate information on effects. In acute studies the requirement for groups and number of animals in groups is related to the reliable determination of acute toxic effects and the estimation of a median lethal dose. In subchronic and chronic testing the numbers are related to the detection of effects, providing sufficient animals for an acceptable investigation of toxic mechanisms and giving an indication of a "no effect level".

A sufficient number of animals should be used so that at the end of the study enough animals in every group are available for thorough biological evaluation. After considerable discussion, it was agreed that for rodents each dose group and concurrent control group should contain at least 10 animals of each sex. For non-rodents, a minimum of four animals of each sex is recommended.

Carcinogenesis bioassays are tools of relatively low sensitivity because of limitations imposed by both experimental conditions and resources. Positive results may be obtained in tests with the use of a small number of animals if the test is otherwise adequately designed and conducted and the tumour response is significant. To support a negative conclusion, however, a larger number of animals is generally used in a carcinogenesis bioassay over that used in other toxicity tests, so that at the end of the study enough animals in every group are available for thorough biological and statistical evaluation.

The approach to the use of control groups is conventional with untreated control groups in subchronic, skin sensitisation and teratogenicity studies (vehicle control groups where required) and periodic positive controls in skin sensitisation studies.

Limit Testing

Many chemicals will only be toxic to man under relatively extreme conditions. In such cases investigation of toxicity with multiple dose groups becomes an academic exercise in which the bulk of the test chemical becomes the limiting factor. Appropriate short-term test

guidelines include recommendations for limit testing, in which one dose of suitable magnitude will serve to delineate the presence or absence of a toxic hazard. These limit values are not presented as absolute values, but rather as a basis on which reasoned toxicological assessments of risk can be made.

Structure-Activity Relationships

There is a growing background of knowledge on the relationships between chemical structure, physical properties and toxicity. The developments are exciting and offer a prospect of an increasing predictive capability in assessing the toxicity of chemicals, particularly those in certain groups. However, these developments are at an early stage and, at the present time, evaluation of the toxicity of chemicals on the basis of structural analogies could give misleading results. The need to study each chemical on an individual basis remains.

ISSUES INVOLVED IN TOXICITY TESTING

The major issues stem from the fact that toxicology is in a stage of rapid development, and harmonization of approaches to testing rests on skill and judgement and not necessarily on purely scientific criteria. On that basis the guidelines represent an agreed basic approach which must serve as a foundation for future development and refinement. Development can only take place as a result of experience, and the Group considers it important that a start is made by first using the guideline methods and then ensuring that their performance is evaluated so that any refinements found necessary can be introduced in an agreed and harmonized manner.

It has already been stressed that the evaluation of toxicity tests and their extrapolation in the evaluation of human health hazard is not straightforward and that techniques are still in the process of development. This is another area in which the Group considers that co-ordination, with possible health monitoring of defined human populations, will be needed to ensure that progress in human health hazard evaluation is sustained.

OECD SHORT TERM AND LONG TERM TOXICOLOGY GROUPS
LIST OF PARTICIPANTS

I. Carruthers*	(Australia)
R. Cumming*	(Australia)
M. Delcour-Firquet*	(Belgium)
A. Lafontaine*	(Belgium)
G.C. Becking	(Canada)
G.S. Wiberg*	(Canada)
I. Knudson*	(Denmark)
E. Poulsen*	(Denmark)
J. Schou*	(Denmark)
I. Chouroulinkov	(France)
G. Smagghe*	(France)
H. Frohberg*	(Germany)
H.P. Gelbke	(Germany)
D. Kayser*	(Germany)
P.J. Kramer*	(Germany)
M. Kunde	(Germany)
W. Lingk*	(Germany)
P. Muhs	(Germany)
H. Stotzer*	(Germany)
H. Zeller*	(Germany)
G. DellaPorta*	(Italy)
D. Misiti*	(Italy)
T. Akimoto*	(Japan)
M. Ikeda	(Japan)
H. Iwasaki*	(Japan)
T. Kamiya*	(Japan)
M. Katoh	(Japan)
Y. Omori*	(Japan)
V.J. Feron*	(Netherlands)
R. Kroes	(Netherlands)
G.J. van Esch*	(Netherlands)
E. van Julsingha*	(Netherlands)
M. van Loghten*	(Netherlands)
E. Dybing	(Norway)
G.S. Dominguez*	(Switzerland)
C. Klotsche	(Switzerland)
K. Schaerer*	(Switzerland)
E. Stenger*	(Switzerland)
K. Fletcher	(United Kingdom)
E.M.B. Smith	(United Kingdom)
Chairman	
Short Term Toxicology Group	
J.R. Beall	(United States)
M.E. Brown*	(United States)
W.M. Butler*	(United States)
W. d'Aguanno*	(United States)
Chairman (Initial)	
Long Term Toxicology Group	
S. Green*	(United States)
R. Hehir*	(United States)
R.B. Jaeger	(United States)
C.R. Morris*	(United States)
N.P. Page*	(United States)
Chairman (Final)	
Long Term Toxicology Group	
R.L. Raleigh*	(United States)
A. Berlin	(CEC)
M. Biart*	(CEC)
W.J. Hunter*	(CEC)
K. Krisor*	(CEC)
M.T. van der Venne*	(CEC)
A.P. Walker**	(OECD)

Observers:

C. Agthe*	(WHO)
J. Parizek*	(WHO)
G. Vettorazzi*	(WHO)

* Attended only part of the meetings
** Adviser to OECD