Recognition of, and Differentiation between, Adverse and Non-adverse Effects in Toxicology Studies

Technical Report No. 85

ISSN-0773-6347-85

Brussels, December 2002

ECETOC Technical Report 85

European Centre for Ecotoxicology and Toxicology of Chemicals
4 Avenue E. Van Nieuwenhuyse (Bte 6), B-1160 Brussels, Belgium.

Adverse and Non-Adverse Effects in Toxicology Studies

CONTENTS

SUMMARY 1

1. INTRODUCTION 2

1.1 Background 2
1.2 Scope of report 3

2. DEFINITIONS AND APPROACHES 5

2.1 Review of existing definitions 5
2.2 Recommendations for a standard set of definitions 5

3. INTERPRETATION OF TOXICOLOGICAL DATA: A STRUCTURED APPROACH 7

3.1 Discriminating factors 'A' for determining whether an effect is treatment related 9
3.2 Discriminating factors 'B' for determining effects that are adverse 19

4. CONSEQUENCES OF NOEL VS NOAEL IN RISK ASSESSMENT 32

5. CONCLUSIONS 33

APPENDIX 1 35

APPENDIX 2 44

BIBLIOGRAPHY 46

MEMBERS OF THE TASK FORCE 49

MEMBERS OF THE SCIENTIFIC COMMITTEE 50

SUMMARY

One of the most important quantitative outputs from toxicity studies is the identification of the highest exposure level (dose or concentration) that does not cause (treatment-related) effects that could be relevant to human health. A review of regulatory and other scientific literature, and of current practices, has revealed a lack of consistency in definition and application of frequently used terms such as 'no observed effect level' (NOEL), 'no observed adverse effect level' (NOAEL), 'adverse effect', 'biologically significant effect' or 'toxicologically significant effect'. Moreover, no coherent criteria were found that could be used to guide consistent interpretation of toxicity studies and, in particular, the recognition of, and differentiation between, adverse and non-adverse effects.

This report addresses these issues, first by proposing a standard set of definitions for the key terms such as NOEL and NOAEL that are frequently used to describe the overall outcome of a toxicity study. Secondly, a structured approach is proposed that will assist the toxicologist in arriving at consistent study interpretation. There are two main steps to the approach. In the first, the toxicologist decides whether differences from control values are treatment-related effects, or occur by chance. In the second step, only those differences judged to be treatment-related effects are evaluated further, in order to discriminate between those that are adverse and those that are not. For each step, criteria are described that form the basis of consistent judgements.

In differentiating an effect from a chance finding, consideration is given *inter alia* to dose response, spurious measurements in individual parameters, the precision of the measurement under evaluation, ranges of natural variation and the overall biological plausibility of the observation. In discriminating between the adverse and the non-adverse effect, consideration is given to whether the effect is an adaptive response, whether it is transient, the magnitude of the effect, its association with effects in other related endpoints, whether it is a precursor to a more significant effect, whether it has an effect on the overall function of the organism, whether it is a specific effect on an organ or organ system or secondary to general toxicity or whether the effect is a predictable consequence of the experimental model.

To arrive at an overall judgement in the interpretation of complex studies, it is important to apply a 'weight of evidence' approach that takes into account the criteria proposed in this report. The use of the structured scheme will contribute to improved consistency of individual study interpretation that is the foundation of reliable prediction of chemical hazard and risk.

1. INTRODUCTION

1.1 Background

Fundamental to the evaluation of the safety of chemicals is an understanding of their intrinsic toxicological properties. Most commonly, such data are derived from toxicological studies that are required by national and international regulatory organisations such as United States Environment Protection Agency (EPA), Japan Ministry of Agriculture Fisheries and Food (Japan MAFF) and the European Commission (EC). In many of these studies, observations are made in laboratory animals exposed to a range of doses or concentrations by the most appropriate route, or routes, of administration. These data are used in two ways. Firstly, to determine the potential hazard by taking into account the route of exposure and nature of any observed adverse effects; secondly, to gain an understanding of the potential risks to humans (under defined conditions) by comparing the dose at which these effects occur with known or estimated human exposures. The critical determining factors will be the adverse effect(s) for which the protection of the exposed population is required and the exposure level(s) at which such effects do and do not occur.

Central to defining both hazard and risk is a clear understanding of whether the treatment-related changes observed in the studies constitute an adverse effect on the test species, or can be considered non-adverse e.g. a healthy, adaptive, response. The growing complexity of guideline studies, and the increasing number of measurements required therein, present major challenges to the reliable identification of adverse effects of chemicals.

A Task Force was thus commissioned to provide guidance on the appropriate evaluation of adverse versus non-adverse effects in toxicity studies. The Terms of Reference were as follows:

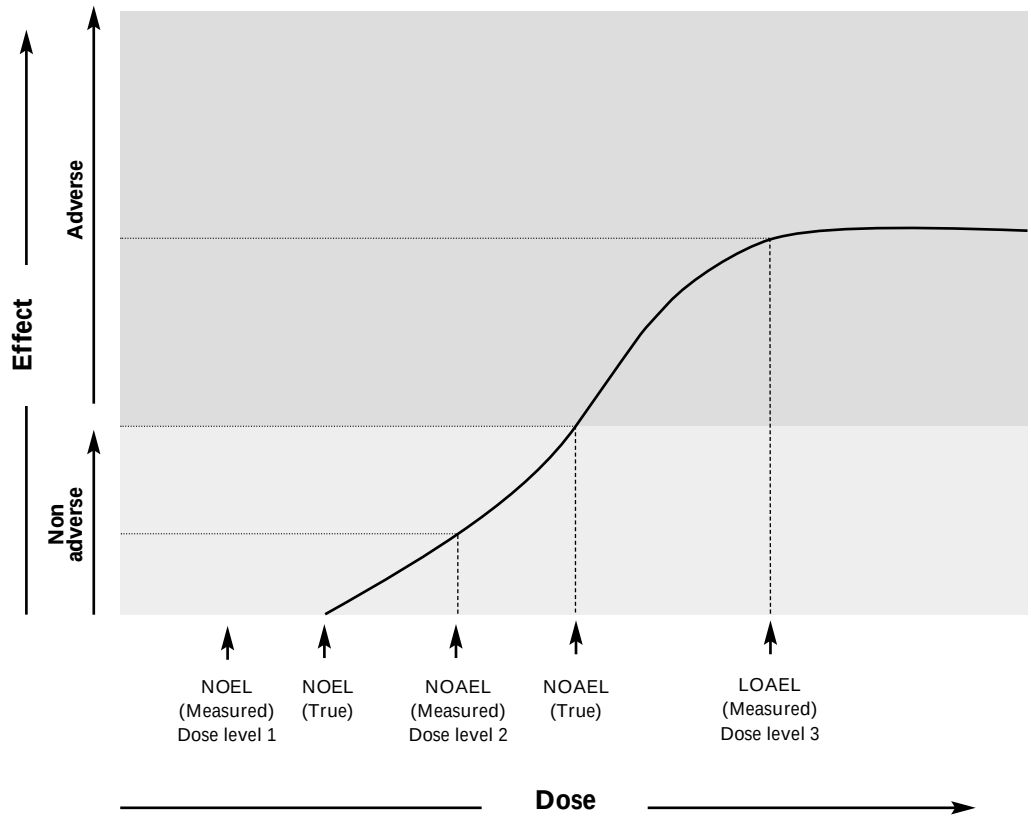
- In the context of the hazard assessment of chemicals, review the current approaches to the uses and definitions of, and criteria for, adverse versus non-adverse effects (addressing such aspects as statistical and biological significance);
- identify, and illustrate with examples, the problems presented by the above definitions, uses and criteria and make recommendations for their resolution;
- give guidance on the implications for risk assessment of a study with a no observed adverse effect level (NOAEL) as opposed to a no observed effect level (NOEL);
- give guidance on a process for evaluating data towards establishing NOAELs and NOELs.

1.2 Scope of report

Observational and descriptive toxicology is able to provide a detailed, qualitative characterisation of a given change. However, to attain the goal of predicting safe conditions for exposure of humans, it is necessary to describe an observed toxicity in terms of its limits i.e. the lowest dose at which an adverse effect is first observed or the highest dose at which it is absent. It follows that a dependable and scientifically robust approach to safety assessment requires consistency in defining those properties / characteristics of an effect, which make that effect adverse. Equally important is the ability to define clearly why some changes in the treated animal(s) can be considered not to be adverse.

It should be appreciated that toxicity studies are of necessity limited to a small number of quantitative observation points (dose or concentration levels), although the biological response may represent a continuum of change with changing dose. For this reason the outputs of such tests in terms of quantitative indices (e.g. adverse-effect and no-adverse-effect levels), and in the derived shape of the dose response curves, represent only an approximation to a true description of the biological effect under study (shown graphically in Figure 1).

Figure1: Quantitative outcome of toxicity studies is dependent on observation points



This document begins with a critical review of existing definitions for terms such as 'effects', 'adverse effects' and 'NOAELs' in order to (1) identify areas of inconsistency and (2) make recommendations as to how these inconsistencies could be resolved. The scope is intended primarily to encompass those toxicology studies designed ultimately for human health risk assessment that identify (or are capable of identifying), adverse effects and the dose levels at which those effects occur or are absent. Screening tests (e.g. *in vitro* tests, short-term tests for oncogenicity potential and those being developed for endocrine active substances), are considered to be outside of the scope of this initiative. The purpose of such tests is to prioritise future action and not to define the dose level at which no adverse effects occur.

Although the guidance provided on study interpretation is intended to apply within the context of a single study, it is recognised that in the overall evaluation of the hazard associated with a test substance, a weight of evidence approach should be taken. Typically, this would involve consideration of the outcome of other studies with the same chemical, other related investigations with analogous chemicals and / or knowledge of effects in other species. Similarly, the guidance is confined to interpretation of effects observed within the defined experimental animal test systems; the relevance of the endpoints evaluated and their appropriate extrapolation in assessing hazard to human health is outside the scope of this report.

This document is intended as a reference for industrial, academic and regulatory toxicologists, as well as those individuals less familiar with the evaluation process, to aid the formulation and application of consistent judgements in the interpretation of toxicity data. This initiative is one of three component parts of a coherent programme aimed at providing guidance on the process of human health risk assessment of chemical substances (ECETOC, 2003a, b).

2. DEFINITIONS AND APPROACHES

2.1 Review of existing definitions

A review of existing definitions was conducted to:

- Identify any existing, accepted, standard definitions;
- examine the level of consistency between organisations / authorities and the level of guidance provided;
- highlight those cases where definition is needed.

The review confirmed / revealed:

- There are no consistent standard definitions for the terms NOEL, or NOAEL or the corresponding terms 'lowest observed effect level' (LOEL) and 'lowest observed adverse effect level' (LOAEL).
- Level of guidance on differentiating 'adverse' effects from 'non-adverse' effects is variable. Most definitions are not accompanied by guidance or by criteria.
- Terms 'toxicity' and 'adverse' are often used interchangeably.
- There is a need for definitions of 'adverse' and 'biologically significant'.
- There is a need for an overall structure for interpreting the effects observed in toxicity studies.

From the review of the above criteria and definitions, it is apparent that to resolve the confusion, a form of words that describes more clearly the terms NOEL and / or NOAEL is needed, along with separate definitions of the phrases 'adverse effect' and 'biologically significant'. Agreement on these definitions is fundamental to development of coherent criteria that can be used to differentiate adverse from non-adverse effects.

2.2 Recommendations for a standard set of definitions

In the context of hazard identification, the NOAEL is the only meaningful expression that describes the highest experimental point (dose) that is without adverse effect. However, the concept and use of the NOEL permeates the literature and regulatory processes. In many cases (see Appendix 2), the terms NOEL and NOAEL are used interchangeably and / or the definition of NOEL uses the concept of 'no adverse effects'. Thus it is important to note that many organisations will actually be employing NOAELs in their regulatory process, even though they may refer to them as NOELs. In the absence of an adequate standard definition for NOEL, the Task Force proposed the following definition:

NOEL - *The highest exposure level at which there are no effects (adverse or non-adverse) observed in the exposed population, when compared with its appropriate control.*

While the Task Force considers this definition of NOEL to be clear and correct, the use of this term is considered to be too simplistic, since it does not discriminate between effects that are adverse from those which are not adverse. For this reason, the Task Force recommends that NOAELs be used for the purpose of developing soundly based hazard assessments of chemical substances. The guidance described in this report for evaluation of toxicity data is consistent with the following recommended definitions. It is important to note, as explained below, that the definition relies on initially differentiating effects from 'changes' or 'differences' from controls.

The EPA (1995a) definitions, augmented with the definitions of Chang *et al* (1982) (cited in full in Appendix 2) were used by the Task Force as a basis for the following recommended standard definitions:

NOAEL - *The highest exposure level at which there are no statistically or biologically significant increases in the frequency or severity of adverse effects between the exposed population and its appropriate control. Some effects may be produced at this level, but they are not considered to be adverse or precursors to adverse effects.*

LOAEL - *The lowest exposure level at which there are statistically or biologically significant increases in the frequency or severity of adverse effects between the exposed population and its appropriate control.*

Adverse effect - *A biochemical, behavioural, morphological or physiological change (in response to a stimulus) that either singly or in combination adversely affects the performance of the whole organism or reduces the organism's ability to respond to an additional environmental challenge. In contrast to adverse effects, non-adverse effects can be defined as those biological effects that do not cause biochemical, behavioural, morphological or physiological changes that affect the general well-being, growth, development or life span of an animal.*

Biologically significant effect - *A response (to a stimulus) in an organism or other biological system that is considered to have substantial or noteworthy effect (positive or negative) on the well-being of the biological system. The concept is to be distinguished from statistically significant effects or changes, which may or may not be meaningful to the general state of health of the system.*

In general, there are two types of significant biological responses. Firstly, there are the normal biological responses, which will manifest in response to stress e.g. sweating in exercise, loss of weight when starved. These changes often represent normal homeostatic reactions to stimuli. Secondly, there are the abnormal biological responses, which may be caused by chemicals or other stresses, e.g. blood dyscrasia, hepatotoxicity, renal toxicity, tumours. Either of these types of biological response could be significantly different from the normal baseline when subjected to statistical analysis, but obviously, the latter is of more concern to toxicologists. Therefore, one must be cautious in relating a statistical finding to a true adverse biological effect.

3. INTERPRETATION OF TOXICOLOGICAL DATA: A STRUCTURED APPROACH

It is generally recognised that evaluating the outcome of complex multi-endpoint toxicology studies is not a straightforward exercise. A comprehensive assessment of toxicological data will involve:

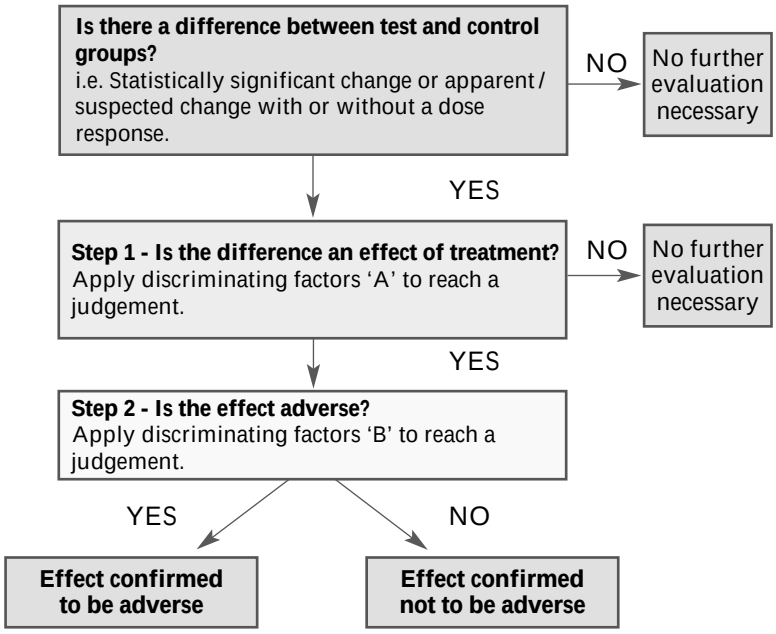
- Expert opinion and judgement, where experience is required to integrate complex and diverse information into a coherent interpretation.
- Recognition that effects may represent a continuum, a threshold or an all-or-nothing response.
- Recognition that in hazard characterisation there are often areas open to interpretation, where description of the outcome in terms of weight of evidence and overall level of concern may be more appropriate and informative than simply commenting on whether an effect is considered to be adverse or not.

In formulating its guidance for distinguishing adverse from non-adverse effects the Task Force considered that the evaluation process would benefit from greater objectivity and consistency through application of a structured approach.

A scheme is proposed in Figure 1 in which the output from toxicological studies is subject to appraisal in a stepwise fashion.

It is emphasised that the evaluation process has been simplified by focusing on individual studies in isolation. This serves the main purpose of giving guidance in establishing which of the differences from control values are treatment-related effects, and of these, which effects should be considered to be adverse. In practice, however, hazard evaluation frequently involves the consideration of data from a collection of studies from different species and study types, culminating in a comprehensive assessment reflecting the overall weight of evidence. At the onset of the evaluation it should be established that there is a difference between the treated and control groups. The evaluation process shown in Figure 1 then comprises two further steps; at each of these steps discriminating factors are proposed which should be considered when making a judgement.

Figure 1: Structured approach to evaluating the outcome of toxicology studies



Step 1 - Is the difference an effect of treatment?

Discriminating factors 'A' are used to differentiate a difference from control values that has arisen by chance from one that is a treatment-related effect. A difference is less likely to be an effect of treatment if:

- There is no obvious dose response;
- it is due to finding(s) in one or more animals which could be considered 'outlier(s)';
- measurement of the endpoint under evaluation is inherently imprecise;
- it is within normal biological variation (i.e. within the range of historical control values or other reference values);
- there is a lack of biological plausibility (i.e. inconsistent with class effects, mode of action, or what is otherwise known or expected of the test substance).

Step 2 - Is the treatment-related effect adverse?

Discriminating factors 'B' are used to differentiate a non-adverse effect of treatment from an adverse effect. An effect is less likely to be adverse if:

- There is no alteration in the general function of the test organism or of the organ / tissue affected;
- it is secondary to other adverse effect(s);
- it is an adaptive response;
- it is transient;
- severity is limited e.g. below thresholds of concern;
- effect is isolated or independent, i.e. changes in other parameters usually associated with the effect of concern are not observed;
- effect is not a precursor, i.e. the effect is not part of a continuum of changes known to progress with time to an established adverse effect;
- it is a consequence of the experimental model.

The appropriate use of the discriminating factors 'A' and 'B' is described and illustrated in greater detail below. Where examples are used, these are either abstracted from the results of specific toxicology studies, or are representative illustrations constructed to demonstrate a point and are derived from the combined experience of the Task Force members.

3.1: Discriminating factors 'A' for determining whether an effect is treatment related

The factors listed under Step 1 (page 8), are used to distinguish differences from control values that are not effects of treatment, from those that are considered to have arisen from treatment with a test substance. In practice, it is likely that a weight of evidence approach will be applied to evaluate the outcome of a study (i.e. taking a combination of these factors into account in order to reach a judgement in differentiating true effects of treatment from other differences from the control population).

For most toxicity studies, the only valid comparisons within a study, are those between treated and concurrent control groups, at each time point during the course of a study. However, when pre-test (baseline) or other previous data are available, as is often the case for toxicity studies in non-rodents, evaluation of differences from these may be the most relevant comparison for evaluation of potential treatment-related effects. This is usually the case for short-term tests, and for indices in long-term tests that show little or no change with time e.g. some serum chemistry measurements. Irrespective of the findings being evaluated, the approach to determining which effects are adverse remains the same.

A-1: A difference is less likely to be an effect of treatment if there is no obvious dose response

The use of dose response is particularly helpful when evaluating the occurrence of statistically significant differences in low and mid-dose groups for routine measurements such as bodyweight, feed consumption, clinical pathology and organ weights. Statistical evaluation of inherently variable biological data can frequently result in chance statistically significant differences between treated and control groups. In these circumstances, it is generally accepted that lack of dose response is an adequate argument for determination that the differences are not related to treatment with the test substance and are, therefore, not effects.

Example:

In a 90-day dietary study in the rat, the following clinical pathology data were obtained for female animals (Table 1). Although statistically significant differences from the control were observed, these were not dose related and thus were considered not to have resulted from treatment. Furthermore, they occurred only at a single dose level, were not associated one with another and were seen in females only (illustrating the weight of evidence approach combining the above-mentioned discriminating factors). The differences from control were considered to have arisen as part of random variation and thus are not 'effects'.

Table 1: Clinical chemistry values for female rats

Parameter	Control	Low dose	Mid dose	High dose
Number examined	10	10	10	10
Plasma glucose (mmol/l)	16.6 ± 2.9	12.2 ± 2.5**	15.7 ± 7.9	15.9 ± 4.0
Plasma total bilirubin (μmol/l)	3.60 ± 0.64	3.69 ± 0.86	4.27 ± 0.82*	3.88 ± 0.47

Values are mean ± standard deviation

*$p \leq 0.05$, ** $p \leq 0.01$, Student's t-test (2-sided)

There are, however, situations that may not lend themselves to dose-response evaluations:

- Changes restricted to the high-dose group. In such cases, where other 'A' discriminating factors do not apply, the changes deserve evaluation as an effect.
- Changes in endpoints observed at lower doses that are masked by overt toxicity (e.g. lethality) at the higher dose. In this case, it is not possible to evaluate a full dose-response relationship.
- Specific findings in mid- or low-dose groups. In some studies such as carcinogenicity and neurotoxicity studies, differences may be observed in mid- and/or low-dose groups which are not present in the high-dose group. These circumstances arise due to different mechanisms of action that are dose dependent (e.g. neuro-stimulatory responses at lower doses and neuro-inhibitory effects at higher doses). Such differences should be considered as effects, although there is no continuous dose-response relationship.

A-2: A difference is less likely to be an effect of treatment if it is due to finding(s) in one or more animals that could be considered outlier(s)

Outliers are extreme individual findings (high or low) that are widely divergent from the main body of a group of data and from historical control values. They are observations and/or measurements that are considered incorrect or directly related to other causes (e.g. disease states), as judged by independent observations or other prior information. For example, the outlier may be due to an unobserved technical error. Outlying values can be detected by visual inspection of the data, use of a scattergram, or appropriate statistical methods (Hamada *et al*, 1998). Differences identifiable only on the inclusion of outlying values are not considered to be effects. In some instances, extreme values can be associated with a specific clinical disease; an example of this would be an animal with a liver tumour, where evidence of altered liver function is encountered in evaluation of clinical chemistry parameters. This does not of itself imply an effect on the whole group and the mean should be recalculated with omission of the extreme values to determine whether an underlying trend remains. As a further example, understanding the influence of extreme values is important in reproduction and development studies. In these studies, the litter is the appropriate unit of evaluation and it is important to understand that changes are often dependent on litter size. An example is foetal or pup weight. Within limits, the total litter weight is relatively constant, but at the extremes of litter size this is no longer true and a few small or large litters may bias mean weights

in a manner that is unrelated to administration of the test substance. In such circumstances, it may be necessary to examine the distribution of litter size and to compare test and control litters of similar size to determine whether the effect is real. Nevertheless, the possibility should not be discounted that a small number of animals may be more sensitive to compound-related effects.

A-3: A difference is less likely to be an effect of treatment if the measurement of endpoint under evaluation is inherently imprecise

Evaluation of differences between control and treated values should include a review of the inherent precision of the measuring methods. Imprecision may result from human and / or instrument factors such as reproducibility, technology and bias. There is, therefore, opportunity for differences between treatment groups and control groups to occur simply based on the precision of the measurement / observation. In addition, the nature of some data (e.g. quantitative data with a normal value near zero) provides for small changes to be statistically different from control when, in reality, they are indistinguishable one from another.

Example:

Table 2 displays a data set for anogenital distance (AGD) from the F2 generation of a multi-generation dietary study in rats (Tyl *et al*, 1999). The apparent increase in AGD can be challenged as an effect of treatment for several reasons (see also Section A-5 on biological plausibility) including whether the measurements for the experimental groups are distinguishable from control based solely on the precision of the measurement.

Table 2: Anogenital distance (AGD) for female rats

Parameter	Control	Low dose	Mid dose (1)	Mid dose (2)	High dose
Number examined	26	26	29	30	27
AGD (mm)	0.76 ± 0.02	0.79 ± 0.01*	0.81 ± 0.02*	0.84 ± 0.02**	0.79 ± 0.01*
Range	0.6 – 1.5	0.6 – 1.25	0.6 – 1.5	0.6 – 1.25	0.6 – 1.25

Values are mean ± standard deviation

*$p \leq 0.05$; **$p \leq 0.01$ Analysed using Kruskal-Wallis test to determine if significant differences were present among the groups, followed by the Mann-Whitney U test for pair-wise comparisons to the control group, if the Kruskal-Wallis test was significant. In addition, Jonckheere's test for k independent samples was used to identify significant dose-response trends (test not indicated on table)

The only absolute values recorded in this study for individual animals were 0.60, 0.75, 0.80, 1.00, 1.25, or 1.50 mm. Thus, rounding the mean is considered justifiable, leading to a value of 0.8 mm for each group. Based on the precision of the measurement alone, the minor deviations between the mean values cannot be considered effects, although statistical significance was obtained. Supporting reasons for lack of association with treatment (e.g. no dose response or biological plausibility), further substantiate the conclusion that the difference in AGD is not a treatment-related effect.

A-4: A difference is less likely to be an effect of treatment if the difference is within normal biological variation (i.e. within the range of historical control values or other reference values)

The use of historical control data should be viewed as a tool for developing a better understanding of the events or apparent differences observed within a study. They should not be seen only as a convenient device for discounting unwanted or 'difficult' findings.

The purpose of a concurrent control group is to represent the normal spectrum of untreated values. Inevitably, the selection process used in assigning animals to experimental groups and the group size in toxicity studies will result in the concurrent control representing only an approximation to the entire control population. Use of a wider data set of control values is often of benefit as it may provide a more appropriate demonstration of the true mean for the population and the variability of that mean.

Historical control data may be used in three primary ways:

1. Identification of aberrant control values

In order to identify aberrant control values, there is a need to understand if the concurrent control group is consistent with the larger population of controls or if it is atypical. When it is determined that the value for a concurrent control group is atypical, individual animal data should be examined to determine if high or low values fall outside the historical range and are a source of bias in calculating the mean. If such outliers are excluded (or none exist), the value for the concurrent control is outside or at the extreme of the historical range and the mean for the treated group is within the historical range, then the difference in the treated groups may be considered unrelated to treatment and not an effect. However, it is important first to take account of any drift in the historical data or the procedural conditions of the study in order to explain this deviation; if this can be ruled out, the control value may be considered to be outside the normal range for that parameter.

Example:

Table 3 shows testes weight data reported in a 4-week inhalation study in rats. The absolute weights of the testes were statistically significantly decreased in all treatment groups except the high dose group.

Table 3: Absolute testes weights in rats

Parameter	Control	Low dose	Mid dose (1)	Mid dose (2)	High dose	HC
Mean absolute testes weight (g)	3.42 ± 0.18	3.02 ± 0.29*	3.06 ± 0.05**	3.10 ± 0.20*	3.17 ± 0.17	3.19 ± 0.15
% of concurrent control	100	88.4	89.4	90.7	92.7	93.3

Number examined = 5 per group
*p ≤ 0.05; ** p ≤ 0.01, Kruskal-Wallis, and Wilcoxon test (two sided)
HC = Historical control

There was no concentration-response relationship or statistically significant change in the relative weights. No corresponding histopathological findings different from the controls were observed. Furthermore, the mean absolute testes weights of the treated groups were within the range of historical control values of 15 other 4-week inhalation studies (mean ± Standard Deviation: 3.19 ± 0.15 g; minimum 2.88 g; maximum: 3.36 g) performed with the same rat strain, whereas the mean weight in the concurrent control group was even higher than the maximum historical control value. Therefore, the concurrent control group is atypical and the numerical 'decrease' in testes weights in the treated groups was regarded as a consequence of the aberrant concurrent control values, and not a treatment-related effect.

2. Understanding relevance of low-incidence findings

For low incidence data, knowledge of the range of historical values is important in differentiating a genuine effect from a spurious difference from the concurrent control.

Evaluation of low incidence data (e.g. incidence of tumours, some foetal malformations) is frequently aided by the use of historical control data. By the nature of the low incidence, it is possible for a treated group to show a low spontaneous incidence, while the control group has lower or even no incidence. In these cases, the historical control will provide information on the overall spontaneous occurrence of the finding. If the treated group incidence falls within the larger population range, the difference from the concurrent control may be considered not to be an effect.

Example:

Table 4 shows tumour incidence data from a 2-year bioassay in rats. The test substance is a non-genotoxic carcinogen with tumours developing only at cytotoxic dose levels (resulting in individual cell necrosis, cellular degeneration-regeneration and hyperplasia).

Table 4: Non-neoplastic (hyperplastic) and primary neoplastic renal changes in male and female rats

Observations through 24 months	Treatment (males)				Treatment (females)			
	C	L	M	H	C	L	M	H
Number examined	50	50	50	50	50	50	50	50
Hyperplasia, tubular epithelium, multifocal	1	0	1	22*	0	0	0	2
Adenoma, cortical, papillary, basophilic	0	0	0	4T	0	0	0	2
Adenocarcinoma, cortical, basophilic, bilateral, no metastasis	0	0	1	0	0	0	0	0
Adenocarcinoma, cortical, basophilic, unilateral, no metastasis	0	0	0	8*T	0	0	0	4T
Adenocarcinoma, cortical, basophilic, unilateral, metastatic	0	0	0	9*	0	0	0	0
Adenocarcinoma, cortical, basophilic, bilateral, metastatic	0	0	0	1	0	0	0	0
Total with primary neoplasm	0	0	1	20*	0	0	0	6*T

C = Control; L = Low dose; M = Mid dose; H = High dose

*Statistical difference from control, Yates chi-square test, $p \leq 0.05$

T Statistical linear trend, Cochran-Armitage test, $p \leq 0.05$

At the high dose, tumours occurred in 40% of males and 12% of females, a finding that was clearly related to the toxicity of the test substance. A comparable tumour was seen in a single mid-dose male (2%), but no renal neoplasms were observed in controls. The absence of an increased incidence of non-neoplastic effects at the mid-dose, considered a prerequisite for test substance-related neoplasia, suggests the tumour may be spontaneous. The fact that the neoplasm was bilateral and malignant supports this suggestion. However, the background incidence of the tumour in untreated controls provides the most compelling supplementary information. In comparable studies of 24 months' duration, initiated prior to or at the same time as the case study, renal adenocarcinomas in control male rats were reported in 4/19 studies with a frequency of 1.0 to 2.0% (Lang, 1992). The tumour in the mid-dose group male animal in this study was therefore within the historical control range and not considered to be an effect of treatment.

3. Understanding relevance of high-incidence findings

High-incidence findings are, by definition, common occurrences (e.g. the 'variants' observed in external, visceral and skeletal examination in developmental toxicity studies). Use of historical data will give information on what range is considered normal for the species under test. Due to the importance and variety of circumstances for use of historical control data in this context, three examples are provided.

Example 1:

Table 5 shows tumour incidence data from a 2-year study in B6C3F1 mice. The data suggest a dose-related increased incidence of hepatocellular adenoma in males and, to a lesser extent, in females.

Table 5: Hepatocellular tumours in male and female mice

	Treatment (males)						Treatment (females)					
Observations	**C**	**L**	**M**	**H**	**HC**	**CS**	**C**	**L**	**M**	**H**	**HC**	**CS**
Number examined	50	50	50	50	550	50	50	50	50	50	550	50
Adenoma	8	10	15	19*	16 6-25‡	25	6	6	12	10	9 4-12‡	12
Carcinoma	5	8	6	9	8 3-13‡	11	1	3	2	8*	2 0-5‡	4
Combined	13	15	20	27*	21 12-30‡	30	7	9	13	15	11 5-16‡	16

C = Control; L = Low dose; M = Mid dose; H = High dose; HC = Historical control data - mean and range;
CS = Control from contemporaneous study; ‡ = Range
*$p \leq 0.05$, Fisher 2x2

There was also an apparent increased incidence of hepatocellular adenocarcinoma in high-dose females. Both of these tumour types are common in this strain of mice (Carmichael *et al*, 1997). The historical control data show that the incidence in concurrent control groups fell within the reference ranges but with an incidence of adenoma in males and females that was close to the lower end of the historical range. When compared with the historical mean values, dose-related trends and significant inter-group differences disappeared for most of the treated groups. The only apparently unequivocal effect that remains is the increased incidence of carcinoma in high-dose females.

Example 2:

Table 6. Tumour data from a 2-year drinking water study in rats show an increase relative to concurrent control of LGL leukaemia, a common spontaneous finding in rats of this strain and age.

Table 6: Large granular lymphocyte (LGL) leukaemia in male and female F344 rats

Observation	Sex	Control	Low dose	Mid dose	High dose	Next study (same lab.)	HC
Number examined		100	100	100	100	50	NS
LGL incidence in spleen (%)	Male	43	51	40	46	66	32-74 (mean 50-58)
	Female	24	41	41*	53**	44	14-54 (mean 28-38)

*p≤ 0.05, **p≤ 0.01, Fisher's exact test
NS = Not stated

Evaluation of the outcome of the study relative to an understanding of natural occurrence of this common finding reveals that the incidence in all treated groups lies within the range of untreated values. For the females, the concurrent control is at the lower extreme of the historical range. In this case it can be concluded that the differences from control observed in this study are not effects of the test substance.

Example 3:

Table 7 shows incidence data for foetal skeletal findings from a developmental toxicity study by gavage in the New Zealand White rabbit.

Table 7: Skeletal findings (% incidence) in New Zealand White rabbits

Skeletal finding	Historical Control	Control	Low dose	Mid dose	High dose
Odontoid – partially ossified	26.3-45.7	40	64.5**	61.7**	72**
Transverse process of 7th cervical vertebra partially ossified	0-6.7	6.7	0.8*	0.7*	1.0**
Transverse processes of 3rd lumbar vertebra fully ossified	2.9-13.8	8.0	0.8**	1.3*	2.5
27 Pre-sacral vertebrae	14.6-36.5	28.0	58.9**	55.1**	59.5**
Unossified 5th sternebra	2.6-13.1	12.7	3.2**	3.4**	5.2*
Partially ossified 5th sternebra	13.3-52.0	52.0	32.3	28.9**	24.6**
Partially ossified 6th sternebra	0-8.	0 8.	0 7.3	6.7	4.2
13th rib short and floating	4.3-14.0	4.0	5.6*	2.7**	5.9*
13th rib normal length	17.1-55.2	42.0	78.2**	82.6**	81.4**

*p≤ 0.05, ** p ≤ 0.01, Student's t test

The shaded data are those dismissed as effects of treatment as they fall within the range expected for untreated animals.

The data show a number of increased and decreased incidences of skeletal findings across all dose groups. In several cases, although statistically significant, the incidence in treated groups clearly falls within the historical range for the specific finding.

As a general point, it is important to emphasise the need for judgement in deciding if a historical data set is appropriate for use in viewing the results of a given study in the context of wider natural variation. Acceptance criteria for the selection and use of historical data have been suggested (Paynter, 1984; Haseman *et al*, 1984), which outline the minimum specification for consideration; studies need to have been conducted in the same strain, age and sex of experimental animal obtained from the same animal supplier, should be from the same conducting laboratory and should be reasonably contemporary to the study under evaluation. Ideally, historical control data should only include studies conducted within an appropriate time period on either side of the study under review, with identification of study methodology (e.g. pre-sampling conditions such as fasting or non-fasting, assay methodology for study parameters, histopathological criteria for lesion identification, and time of terminal sacrifice) that could have affected the results. Literature values for normal ranges that do not specify the method by which they were obtained, should be used with caution.

A-5: A difference is less likely to be an effect of treatment if there is a lack of biological plausibility (i.e. is inconsistent with class effects, mode of action or what is known or expected of the test substance)

In many cases, biological, chemical or physical properties of the test substance, and / or results from testing of the same or similar materials, provide information that can help determine whether a response to the test substance is biologically plausible. For example, in some cases, measurements (not observations) may be excluded from consideration as effects from treatment with the test substance if they are clearly outside the biological / physiological range. These measurements will result from experimental or human error. Examples would include bodyweights higher than maximum attainable for the species / strain, negative feed consumption values, and haematology / clinical chemistry values which are physiologically impossible:

Example:

Table 8 shows data from a subchronic dermal toxicity study in rats, and further illustrate the importance of considering the biological plausibility of experimental observations.

Table 8: Adrenal weights in male and female rats

Observation	Treatment (Males)				Treatment (Females)			
	Control	Low dose	Mid dose	High dose	Control	Low dose	Mid dose	High dose
Number of animals	10	10	10	10	10	10	10	10
Bodyweight (g)	213	209	210	206	130	130	126	129
Absolute adrenal weight (g)	0.050	0.046	0.051	0.051	0.048	0.047	0.049	0.055
Relative adrenal weight (g / 100g bw)	0.023	0.022	0.024	0.025*	0.037	0.036	0.039	0.042*

*$p \leq 0.05$, Student's t test

The data indicate a statistically increased relative adrenal weight for high-dose males and females. Reference to existing data generated for this test substance by the oral route indicated that this was not a biologically plausible treatment-related finding. The test substance was completely absorbed by the oral route and, in most part, excreted un-metabolised. By comparison, absorption through the skin was very limited and systemic exposure by this route was almost two orders of magnitude lower than by the oral route. The adrenal gland was not a target organ by the oral route at the highest dose tested. Therefore, the statistically identified increase in adrenal weight in the dermal study was not related to treatment.

3.2: Discriminating factors 'B' for determining effects that are adverse

The factors considered under Step 2 (page 8) are used to help differentiate between those changes that are effects of treatment and considered non-adverse, from those that are considered to be adverse effects. It is likely that in practice, a weight of evidence approach will be taken i.e. that combinations of these factors will need to be considered in order to reach a judgement in differentiating the non-adverse from the adverse.

B-1: An effect is less likely to be adverse if there is no alteration in the general function of the test organism or of the organ or tissue affected

There may be effects in toxicity studies that do not represent any functional impairment in the test organism. Such effects are considered not to be adverse. This includes inhibition of plasma butyrylcholinesterase and reductions in some standard haematological and clinical chemistry parameters, such as prothrombin time, enzyme activities and concentrations of cholesterol and bilirubin in blood. Some of these examples are considered in more detail to illustrate this point.

Example 1:

Butyrylcholinesterase activity is markedly reduced by anticholinesterase compounds, such as organophosphate and carbamate insecticides. Unlike acetylcholinesterase, which is a neurotransmitter, butyrylcholinesterase has no neurochemical role. Hence plasma butyrylcholinesterase activity may be a useful qualitative indicator of exposure to a substance with anticholinesterase activity, but there is no association between its activity and cholinergic effects and signs of toxicity. Therefore, plasma butyrylcholinesterase inhibition is a toxicologically insignificant finding and is not an adverse effect (WHO PCS, 1998).

Example 2:

Table 9 includes data from a subchronic study in male and female rats.

Table 9: Serum transaminase activities in male and female rats

Parameter	Sex	Control	Low dose	Mid dose	High dose
SGOT [U/l]	Male	110.4±10.26	81.3±7.87	93.5±4.36	68.1±3.95**
	Female	79.5±4.20	81.3±5.21	67.0±4.433	60.9±3.52**
SGPT [U/l]	Male	30.4±4.27	25.3±1.68	19.6±0.9**	16.1±0.83**
	Female	39.7±4.89	34.6±3.58	20.6±1.08**	15.9±0.91**

Number examined = 10 per group

Values are mean ± standard deviation

* $p \leq 0.05$, ** $p \leq 0.01$, Student's t test

Increased activities of certain enzymes in blood, such as alanine and aspartate transaminases, may identify organ (e.g. liver) toxicity. However, decreased activity of these enzymes, as shown above, is generally of no toxicological importance (Willard and Twedt, 1994). Decreased transaminase activities are commonly observed for liver cytochrome P-450 inducers and in long-term studies with reduced food intake and bodyweight gain.

Example 3:

Table 10 shows data on serum cholesterol levels measured in a 1-year study in male dogs.

Table 10: Cholesterol levels (mmol/l) in female dogs in a 1-year study

Observation time	Control	Low dose	Mid dose	High dose
Pretest	2.90 ± 0.3	3.74 ± 0.6	3.01 ± 0.6	2.49 ± 0.5
Week 13	3.39 ± 0.4	3.83 ± 0.8	3.53 ± 0.3	1.63 ± 0.4 **
Week 26	4.36 ± 0.2	4.21 ± 0.4	4.55 ± 1.1	2.02 ± 0.4 **
Week 52	4.66 ± 0.7	4.55 ± 0.7	4.75 ± 0.9	2.16 ± 0.5 **

Number examined = 4 per group
Values are mean ± standard deviation
** $p<0.01$, Dunnett test

Usually cholesterol concentrations increase naturally with the age (growth) of the animals. The data indicate that the test substance affected cholesterol levels in the high-dose group compared to the concurrent control at different time points in the study. However, there was no associated liver toxicity, and because of the direction and magnitude of these changes, they have no clinical importance and are not considered adverse.

B-2: An effect is less likely to be adverse if it is an adaptive response

Living organisms have a capacity to respond to environmental variations and stresses, whether physical or chemical, in order to maintain normal function and survival. Physiological processes are regulated by hormonal and enzymatic control systems which operate at the level of the cell, organ or multi-organ systems. Certain effects may be adaptive responses to general chemical exposure and unrelated to inherent toxicity of the test substance as such. These types of effects include liver enzyme induction and limited liver enlargement as a physiological response to the need for increased metabolic activity and adaptation of the respiratory tract to modified requirements of tissues under exposure.

Example:

Burger *et al* (1989) described histological changes in the respiratory tract that may be assessed as adaptive responses:

- Mucous cell hyperplasia may be induced by dehydration of the nasal epithelium or inhalation of aerosols of a variety of chemicals.
- Exposure to aerosols may cause transitional epithelium at the base of the epiglottis in rats to change into squamous epithelium.
- Macrophage accumulation in the lung after exposure to low solubility materials in the absence of any signs of inflammatory reaction is considered to be a physiological sign of enhanced alveolar clearance activity.
- To further enhance alveolar clearance, alveolar epithelium of the lung may be replaced by ciliated epithelium. This process is called "bronchiolisation" and represents the ultimate adaptive response after exposures to high concentrations.

As the findings described above are steps in a continuum of changes which may progress into adverse effects, much scrutiny must be invested in their examination and interpretation.

B-3: An effect is less likely to be adverse if it is transient

An effect is less likely to be adverse if it disappears during the course of treatment. Such transient effects are common when they result from non-specific responses to treatment such as non-palatability of diets immediately following initiation of treatment or stress from inhalation exposures, gavage dosing, or skin applications / wrapping. When no chemical-specific toxicity occurs, it is possible for effects (e.g. on bodyweights and clinical signs) to disappear following a brief acclimation period.

There is need to make distinction between transient effects which disappear during exposure and effects which recover after exposure ceases. In the latter case, the effect may be adverse during exposure and at least in cases where continuous exposure exists, the effect must be considered in the hazard assessment of the test substance.

Example:

The data shown in Table 11 consider bodyweight and food consumption data from a feeding study in male mice.

Table 11: Bodyweights and food consumption for male mice

Bodyweight (g)				
Week	**Control**	**Low dose**	**Mid dose**	**High dose**
1	22.3 (±2.6)	22.8 (±3.5)	23.2 (±2.9)	22.2 (±3.0)
2	29.9 (±3.1)	29.8 (±3.7)	30.4(±3.1)	24.0 (±3.0)**
3	35.2 (±3.5)	35.0 (±4.2)	36.0(±3.5)	29.0 (±3.4)**
4	37.9 (±4.4)	37.4 (±3.8)	38.8(±3.9)	36.8 (±3.1)
5	39.3 (±4.4)	38.9 (±5.2)	40.3(±4.3)	38.9(±3.6)
6	39.4 (±4.6)	39.1 (±4.9)	40.3(±4.9)	39.2(±4.1)
7	40.9 (±4.5)	40.5 (±5.2)	41.7(±4.8)	40.0(±3.4)
8	41.9 (±4.3)	41.8 (±5.2)	42.9(±4.6)	42.1(±3.6)

Food consumption (g / animal / day)				
Week	**Control**	**Low dose**	**Mid dose**	**High dose**
1	7.0 (+0.5)	7.5 (+0.5)	7.5 (+0.5)	3.9 (+0.9**
2	7.8 (+0.7)	7.8 (+0.4)	8.1(+0.6)	6.9(+0.7)*
3	8.2 (+0.4)	8.3 (+0.4)	8.6(+0.7)	7.5 (+0.9)
4	7.1 (+0.4)	7.0 (+0.4)	7.2(+0.8)	7.5 (+0.6)
5	7.0 (+0.5)	7.0 (+0.4)	7.1(+0.9)	7.1 (+0.7)
6	7.2 (+0.6)	7.4 (+0.4)	7.4(+0.8)	7. 0(+0.8)
7	7.0 (+0.6)	7.0 (+0.3)	7.1(+0.4)	7.5 (+0.6)

Number examined = 12 per group

Values are mean ± standard deviation

** $p \leq 0.01$, * $p \leq 0.05$, Student's t test, two sided

The data show transient effects on bodyweight correlated with changes in food consumption. These changes were attributed to an alteration of taste or odour of the feed leading to initial reduction in food intake accompanied by reduced growth (effects disappear with adaptation of animals). Therefore, these effects of the test substance are not considered to be adverse.

B-4: An effect is less likely to be adverse if the severity is limited e.g. below thresholds of concern

To distinguish between adverse and non-adverse findings the consideration of severity plays an important role. For certain effects a clear demarcation for non-adverse and adverse is definable (see following example). In contrast, for target organ effects which can vary greatly in severity, a continuum of findings is sometimes described ranging from minimal changes to pronounced adverse effects. According to the degree of severity, effects can be categorised as adverse or non-adverse. Non-adverse effects are usually adaptive or compensatory responses or findings which are below a threshold level or do not fulfil the criteria which define their biological significance. However, often the distinction between adverse and non-adverse effects is not clearly defined and interpretation needs scientific judgement on a case-by-case basis.

Example:

In defining a biologically significant depression of brain and erythrocyte cholinesterase, thresholds (such as 20% inhibition) are used. A statistically significant reduction in these enzyme activities by more than this threshold is considered to represent an adverse effect. A statistically significant inhibition of less than this requires more detailed analysis, and interpretation should be made on a case-by-case basis. For this purpose the shape and slope of the dose-response curve, assay variability and correlation with clinical signs should be considered (WHO PCS, 1998).

B-5: An effect is less likely to be adverse if it is isolated or independent

An effect of treatment is less likely to be adverse if the response occurs in isolation or is not associated with effects in other related endpoints. Individual responses may be encountered in large-animal studies, since fewer animals and less homogeneous populations are employed. Careful consideration of a unique response is required. Before discounting an effect in an individual animal, the influence of such factors as the type of effect, the magnitude of response, the relationship of the effect to known toxicity, mechanisms of action and kinetics of the test substance should be gauged. The effect may be considered non-adverse if the response can be attributed to sensitivity of the individual to treatment (i.e. non-specific response to the test substance), or the magnitude of the response is minimal, or the effect is inconsistent with known response to the test substance.

Although less frequent when using homogeneous rodent populations, in some circumstances (e.g. neurotoxicity evaluations) isolated effects may occur without corresponding effects in related measurements. These types of responses are more frequently encountered when large numbers of measurements are made, thus providing opportunity for differences from control due to chance alone to appear to be treatment-related. Knowledge of the affected system is critical for interpretation of these situations and evaluation of the type and magnitude of the effect is important. When an effect occurs in a system in which other related measurements should clearly be affected as well, the effect may be considered non-adverse.

Example:

The data shown in Table 12 were collected in a 13-week inhalation study in female rats to evaluate neurotoxicity.

Table 12: Functional observational battery in female rats

Observation	Control	Low dose	Mid dose	High dose
Number examined	15	10	10	15
Body position Lying on side	1 (7)	1 (10)	3 (30)	0 (0)
Sitting / standing	12 (80)	8 (80)	5 (50)	14 (93)
Rearing	4 (27)	5 (50)	3 (30)	13 (87)**

**Statistically significant ($p \leq 0.01$) by Fisher's exact test
Percentage of animals affected is shown in brackets

The data indicate a statistically significant increase in 'rearing' at the high-dose level. However, 'body position' is one of 42 measurements made for the functional observational battery (FOB) and the only statistically significant difference observed in these animals was for 'rearing' in the high-dose group. Due to the lack of effects on other FOB measurements that would be expected in association with a treatment-related change in 'rearing', this effect in isolation is not considered to be adverse.

B-6: An effect is less likely to be adverse if it is not a precursor to a known adverse effect

An effect may be considered non-adverse if the response has been shown not to progress to adverse toxicity.

The most comprehensive information of this type of process derives from the studies of liver weight effects and their relationship to chronic toxicity and carcinogenicity. The interpretation of increased liver weight should be done on a case-by-case basis taking into consideration (when available) the data obtained in long term studies (carcinogenicity). The change in liver weight should be evaluated under one of the following scenarios:

- Increased liver weight with no evidence of pathological changes:

  Increased liver weight in the absence of adverse histological changes should be considered as an adaptive response (i.e. not adverse) provided data are available to indicate a plausible mechanism e.g. enzyme induction or peroxisome proliferation. In the absence of such explanatory data, a statistically significant increase in liver weight should be regarded as an early marker of a potential adverse effect subject to consideration of the magnitude of the increase.

- Increased liver weight with evidence of hypertrophy:

  Zonal, usually centrilobular, hypertrophy of the hepatocytes is generally regarded as an adaptive effect associated with enzyme induction or smooth endoplasmic reticulum proliferation. In these cases, the magnitude of the increase in the liver weight should be taken into consideration when deciding on the adversity of these effects.

- Increased liver weight with pathological changes but no evidence of tumour induction:

  If increased liver weight and evidence of hepatotoxicity are both present, this should be considered as an adverse effect. The next lower dose level should be considered as the no observed adverse effect level for liver effect.

- Increased liver weight with evidence of tumour induction:

  It has been recognised with a number of carcinogens that there is frequently an association between persistent liver enlargement and the induction of hepatic tumours and that a clear dose-effect relationship is evidenced between the liver weight, clinical pathology indicators of liver toxicity, histopathological lesions and tumour induction. If the increased liver weight is present at the same dose levels as overt hepatotoxicity and tumours, this is clearly an adverse effect. In cases where increased liver weight is present at a lower dose and where there is no detectable hepatotoxicity, the increased weight should be considered as an early indication of the frank toxicity observed at higher dose levels. Therefore, in this situation the increased liver weight is considered to be an adverse effect.

Many other inter-related factors such as the magnitude of liver weight increase, evidence of reversibility, incidence and severity of histopathological changes, evidence of a plausible mechanism, evidence of other adverse effects in other tissues, should also be considered in establishing whether an increased liver weight is regarded as an early indicator of liver injury or as an adaptive response with no evidence of adversity.

Example:

Table 13 shows data from a 90-day toxicity study in rats with a polymeric material admixed in the diet of male and female Fischer 344 rats.

Table 13: Microgranuloma in mesenteric lymph nodes of male and female rats

Severity	**Control**	**Low dose**	**Mid dose**	**High dose**
Minimal	2	5	4	4
Slight	1	0	1	3
Moderate	0	1	0	2
Severe	0	0	0	0
Severity	**Control**	**Low dose**	**Mid dose**	**High dose**
Minimal	4	2	0	3
Slight	7	1	5	3
Moderate	1	10	9	6
Severe	0	0	0	1

Number examined = 15 per group

The data indicate an increased incidence and/or severity of microgranulomas at all three dose levels (dietary concentrations ranged from 1-5%). No other treatment-related effects were observed in the study.

The findings were consistent with similar observations which have been reported in a number of subchronic and chronic toxicity studies in Fischer 344 rats (Ward *et al*, 1993; Shoda *et al*, 1997; Fleming *et al*, 1998). The lesions have also been observed with variable incidence (approximately 10-95%) in controls (Ward *et al*, 1993; Firriolo *et al*, 1995; Shoda *et al*, 1997). Microgranulomas of the mesenteric lymph nodes occur in a non dose-related manner in animals fed high (>1%) dietary concentrations of materials such as white oils and waxes (Fleming *et al*, 1998) and are considered to represent a clearance mechanism to remove foreign bodies introduced from the high doses of chemicals in the feed.

Key to determination of the adversity of lesions related to normal mechanisms (such as clearance) is the relationship to adverse effects from long-term exposure. This type of inflammatory lesion was not associated with the development of neoplasia in a chronic study using medium viscosity liquid paraffin (Shoda *et al*, 1997). Therefore, it was concluded that these lesions represented a clearance mechanism for the test substance and were not in themselves adverse.

B-7: An effect is less likely to be adverse if it is secondary to other adverse effects

In many cases, observations made in toxicology studies appear to be adverse effects of treatment when, in fact, they are secondary and directly attributable to other adverse effects. The most common occurrence of this type of apparent adverse effect is in organ weight data. Since many organs tend to change weight in relation to changes in bodyweight, effects on absolute organ weight may be considered non-adverse if the organ-to-bodyweight ratios are comparable. Evaluation of the relationship of such measurements generally relies on statistical power to indicate appropriate correlation.

Example:

Ashby and Lefevre (2000) recently investigated the relationship between bodyweight and sexual development. Specifically, they studied the influence of bodyweight on the day (age) of preputial separation (PPS) in untreated rats of different bodyweight at the day of weaning. This analysis of weight-stratified control data demonstrated a marked dependence of the day (age) of PPS on the initial animal bodyweight. The relationship between age at PPS and bodyweight at PPS was less marked than when PPS was related to initial (weaning) bodyweight. Using this knowledge, the authors concluded that chemically-induced delays in male sexual development could only be identified with confidence when:

- They were not accompanied by treatment-related suppression in bodyweight or in bodyweight gain;
- expected delay in prepuce separation due to bodyweight change was exceeded (this assertion can be made from the observation that, within limits, there appears to be a set animal bodyweight at which sexual maturation occurs, irrespective of age).

From these and other emerging data, it seems likely that animal bodyweight, rate and nature of weight change, and the timing of these effects, can cause a secondary effect on the age of sexual development in both males and females.

Example:

Table 14 shows skeletal data from a developmental toxicity study in rats.

Table 14: Litter/foetal data for the rat

Observations		Control	Low dose	Mid dose	High dose
Mean foetal weight (g)		4.95	4.89	4.91	4.16**
Minor skeletal defects	% foetuses affected	31.7	38.1	26.7	54.6**
	number of litters affected	24/24	23/24	23/24	24/24
Skeletal variants	% foetuses affected	74.5	74.1	78.4	97.3**
	number of litters affected	24/24	24/24	24/24	24/24
Skeletal defects	% of foetuses affected				
- odontoid not ossified		19.5	20.9	28.8	67.3**
- cervical vertebrae					
- centrum not ossified, 2nd		22.2	29.4	39.9**	86.7**
- centrum not ossified, 3rd		8.6	11.1	20.1*	58.3**
- centrum not ossified, 4th		10.4	8.1	14.1	34.3**
- centrum not ossified, 6th		0.9	0.6	2.6	6.2**
- transverse processes of 4th lumbar vertebra fully ossified		7.0	5.2	7.6	1.2**
- calcaneum not ossified		25.5	30.4	31.2	89.8**

All values are means from 24 litters in each experimental group
** $p \leq 0.01$, * $p \leq 0.05$, Student's t test, two sided

The data indicate a delayed ossification accompanying reduced foetal weights i.e. a delay in foetal development rather than a direct effect on bone tissue. At the high dose only, the reduction in foetal weight is accompanied by a reduction in the degree of ossification of the foetal skeleton. Therefore, this change is secondary to the adverse effect on bodyweight rather than a specific effect on bone formation. The reduced ossification of cervical vertebrae 2 and 3 reported in the mid-dose foetuses was not considered an effect of treatment, since there were no correlated effects in other bones (see Section B-6).

Systemic effects secondary to local effects

Systemic effects may be produced as a consequence of local effects rather than as a direct action of a chemical. In many circumstances, such responses may be secondary to the portal of entry (local) effects and not caused by systemic interaction of the test substance following distribution in the body. This is often in the form of irritation to the skin, the gastro-intestinal tract or the respiratory tract.

In certain studies, local and systemic effects may be observed as a consequence of the study design (e.g. the route of exposure). It is not uncommon to observe adverse local effects at the site of exposure, as this is the site exposed to the highest initial concentration of the test substance. Dermal exposure may lead to local irritation or corrosive effects. In inhalation studies, the effects may be seen as local inflammatory reactions in the nasal and bronchial epithelium or in lung tissue. Dietary or gavage exposure may lead to gastric irritation and there may be local effects at injection sites of compounds administered parenterally. Under these circumstances, local effects clearly should be separated from general systemic toxicity. In defining NOAELs, the purpose of the study should be taken into consideration and a differentiation made between local and systemic effects. Correspondingly, a systemic change(s) secondary to the local effect should be dissociated from the true systemic toxicity and two NOAELs established, one for local and one for systemic effect. According to the purpose of the study, the corresponding NOAEL should be used for further hazard characterisation or risk assessment.

Portal of entry effects are of particular interest in inhalation studies, because the respiratory tract, as a major organ system, may be affected by direct interaction with inhaled materials. Local effects are governed by qualitative and quantitative disposition patterns of the inhaled materials within the organ system and their intrinsic physico-chemical and toxic properties (Greim *et al*, 2001). Inhalation of respirable particulate materials with high biopersistence leads to accumulation of particles in the alveolar region of the lung. Consequentially, an activation of alveolar clearance occurs, with recruitment of alveolar macrophages and transport of material to the mucociliar escalator of the bronchi or via interstitialisation to the regional lymph nodes. Inflammation and fibrotic changes may develop, depending on the amounts of material deposited in the lung and the intrinsic toxicity of the inhaled test substance.

The effects of irritant gases or vapours within the respiratory tract are mainly determined by their water solubility and reactivity. During nasal breathing the highly water-soluble formaldehyde, for example, is readily scrubbed from the air in the upper respiratory tract (mainly the anterior nose), which constitutes the target organ at relevant concentration (Morgan, 1997). On the other hand, ozone penetrates easily into the alveolar region of the lung with respective differences in distribution of irritant effects (Miller, 1995). As these compounds react mainly with, or are metabolised in, the tissues of first contact, their systemic availability at concentrations already producing marked local effects is negligible. Thus, portal of entry effects produced by irritant gases usually occur at much lower exposure concentrations than systemic effects.

Example 1:

Rats were treated dermally with the test substance at dose levels of 0, 10, 100 and 1000 mg / kg bw for 28 days. Signs of local dermal irritation (erythema) were noted at 100 and 1000 mg / kg bw / day. Systemic effects consisted of a decrease in bodyweight and an increase in liver weight at 1000 mg / kg. Based on the reduced bodyweight, the NOAELs for systemic effects and for local effects were 100 mg / kg and 10 mg / kg bw / day respectively.

Example 2:

The data from a 28-day aerosol inhalation study in rats are shown in Table 15.

Table 15: Lung toxicity and haematological changes in rodents

	High dose	
Parameters	**Males**	**Females**
Number of animals	5	5
Lethality (towards the end of the study)	4	1
Reduced general condition	3	4
Respiratory sounds after daily exposure	5	2
Bodyweight (% of control)	77	73
Red blood cell count (% of control)	111	125
Blood neutrophil count (% of control)	305	329
Absolute lung weight (% of control)	182	235
Histopathological lung findings		
- Pneumonia	5	5
- Bronchial necrosis	4	4
- Emphysema	4	5

At a high concentration, the lung toxicity of inhaled aerosols led to severe inflammation and development of emphysema in the lung. These changes formed the basis for a reduction in the general condition of the animals, increased erythrocyte counts probably due to hypoxia, and increased blood granulocyte counts due to pneumonia.

At lower exposure concentrations, respiratory tract findings were present in the larynx only and no other signs of toxicity were observed. The systemic effects on general health and haematology were attributed to the severe changes seen in the lung. Lethality, clinical signs and increased red blood cell counts were ascribed to decreased oxygen exchange in the alveolar region; increased neutrophil counts were effected by pneumonia.

The systemic effects are thus considered to be a consequence of the adverse effects on the lung rather than a direct effect of the test substance.

B-8: An effect is less likely to be adverse if it is a consequence of the treatment regimen

An effect is less likely to be adverse if it can be ascribed to a consequence of the treatment regimen rather than of the test substance. To avoid both animal discomfort and confounding effects, extensive efforts are made to design studies such that any effects arising from the treatment regimen are eliminated. However, effects may arise as reactions to stress from treatment (such as restraining in nose-only inhalation studies, the use of bandages in dermal exposure studies), reactions to the physical properties of the test substance (such as odour and taste) or reactions to low-level physical trauma (handling of the animals). These types of direct responses to aspects of the treatment regimen should be separated from portal of entry and localised effects previously described.

Special considerations – reversibility

Reversibility can be an important factor in the holistic interpretation of toxicology studies. Although not one of the discriminating factors that can be used to distinguish differences from effects and adverse from non-adverse effects, a knowledge of reversibility is often used as a key part of the weight of evidence approach to study interpretation.

In assessing the level of concern assigned to a given biological effect, a change which is readily and completely reversible on cessation of treatment is considered to indicate a lower level of concern. It follows that a knowledge of whether or not an effect is reversible may influence significantly the overall interpretation and differentiation of adverse from non-adverse effects.

4. CONSEQUENCES OF NOEL VS NOAEL IN RISK ASSESSMENT

Consequences for the risk assessment of a substance employing the NOAEL as opposed to the NOEL from a study are demonstrated in the following example.

Example:

Dogs were treated for 6 months at dose levels of 0, 0.2, 2, 100 and 500 ppm. The results revealed a dose-dependent inhibition of the plasma butyrylcholinesterase activity in both sexes at 2, 100 and 500 ppm. However no effect was observed on brain cholinesterase activity. Based on these findings, the EPA concluded that the inhibition of the butyrylcholinesterase activity was an adverse effect and defined the NOEL in this study as 0.2 ppm, (equivalent to 0.005 mg / kg bw / day) (EPA, 2000). This conclusion was used by the World Health Organization (WHO) to set a reference dose at 0.00005 mg / kg bw / day, applying a safety factor of 100. In contrast, WHO concluded from the same study results that plasma butyrylcholinesterase inhibition was a toxicologically-insignificant finding (i.e. not an adverse effect) and defined the NOAEL at 100 ppm based on other systemic effects that were observed at the 500 ppm level (Quest, 1990). The consequence of this decision prompted WHO to use a different endpoint in a different study (1 mg / kg bw / day in a 2-generation study) to derive an ADI of 0.01 mg / kg bw / day.

The overall outcome of the different approaches adopted by the two agencies in their interpretation of the study NOEL and NOAEL resulted in ADIs that differ by a factor of 200.

Table 16: Plasma butyrylcholinesterase activities in 6-month dog study

Observations	Time point	Control	0.2 ppm	2 ppm	100 ppm	500 ppm
Number of animals		7	7	7	7	7
Plasma	pretest	102 ± 25	102 ± 23	104 ± 26	108 ± 37	102 ± 15
butyrylcholinesterase	week 4	97 ± 16	88 ± 21	51 ± 10*	23 ± 7*	11 ± 3*
(Klett units) pretest	week 13	94 ± 21	78 ± 14	46 ± 9*	24 ± 6*	12 ± 3*
	week 26	74 ± 10	71 ± 17	47 ± 5*	19 ± 5*	14 ± 4*
Erythrocyte	pretest	97 ± 17	108 ± 26	100 ± 34	85 ± 15	94 ± 25
cholinesterase	week 4	80 ± 10	93 ± 31	89 ± 26	39 ± 18*	21 ± 9*
(Klett units) pretest	week 13	94 ± 8	100 ± 39	92 ± 23	37 ±16*	26 ± 7*
	week 26	106 ± 15	102 ± 33	84 ± 27	27 ± 17*	6 ± 2*
Brain cholinesterase (Klett units)	week 26	169 ± 30	183 ± 36	160 ± 11	187 ± 21	194 ± 56

*$p \leq 0.05$, Lepage test

5. CONCLUSIONS

In evaluating the intrinsic toxicity of a chemical substance, two factors are key. These are knowledge of the nature and significance of any adverse effect, and the dose or exposure level at which this effect is absent or first observed. To gain this information there is a requirement for consistent definition and application of the concept of adversity in biological systems.

Review of existing definitions of terms such as 'adverse' and of approaches to interpretation of toxicology studies (i.e. consistent recognition of adverse effects), revealed considerable inadequacy and confusion. In order to address these, a set of clear uncomplicated definitions was prepared, drawing on the best advice available. These definitions are as follows:

Adverse effect: *A biochemical change, functional impairment, or pathological lesion (in response to a stimulus) that either singly or in combination adversely effects the performance of the whole organism or reduces the organism's ability to respond to an additional environmental challenge. Contrasted to adverse effects, non-adverse effects can be defined as those biological effects which do not cause physical, physiological, behavioural and biochemical changes that affect the general well-being, growth, development or life span of an animal.*

Having established a clear definition of 'adverse', this was then applied to definition of the key outcomes of toxicity studies:

NOAEL – *The highest exposure level at which there are no statistically or biologically significant increases in the frequency or severity of adverse effects between the exposed population and its appropriate control. Some effects may be produced at this level, but they are not considered to be adverse or precursors to adverse effects.*

For toxicological studies not demonstrating a clear NOAEL, the LOAEL will be the critical value.

LOAEL – *The lowest exposure level at which there are statistically or biologically significant increases in the frequency or severity of adverse effects between the exposed population and its appropriate control.*

The use of a generic, structured approach for evaluating the data generated in toxicology studies was considered critically important in achieving consistency in their interpretation. A two-stage process was found by the Task Force to be the most appropriate and logical. Firstly, discrimination is required between those differences from control values that may be observed in treated groups that are effects of treatment, from those that are not effects of treatment. Having, in this way, identified the genuine treatment-related effects, these are then further examined in the second step in order to differentiate those effects that are adverse from those that are not.

A clear process for achieving this outcome has been proposed. At each level of decision-making, a set of discriminating factors has been identified and their application illustrated by examples. Application of the scheme utilising these factors ensures a more consistent outcome in study interpretation. It is recognised that evaluating the outcome of complex multi-endpoint toxicology studies is not a straightforward exercise. A comprehensive assessment of toxicological data requires expert opinion and judgement. In practice it is important that a weight of evidence approach be taken i.e. combinations of the discriminating factors identified will generally need to be taken into account to arrive at an interpretation of a single toxicology study, or indeed of a toxicology database for a given substance.

The output of this report serves to establish a reliable foundation on which the next stages of a coherent risk assessment process depend. These subsequent steps include the evaluation of the relevance to humans of the effects observed in animal studies, typically by examination of mechanism. When relevance to humans has been established, the magnitude of the assessment factors appropriate for protecting human health can be applied. Guidance on these latter two steps is the subject of separate ECETOC reports (ECETOC, 2003 a,b).

APPENDIX 1

Table 17: Summary of definitions and criteria provided by major regulatory authorities and organisations involved in the development of experimental test guidelines

Term	Narrative definition/comments	Reference
NOAEL	Highest dose level where no adverse effects(a dose response is implied).	OECD 407 (1995a)
NOEL/ NOAEL.	Inconsistent approach with no precise No toxic effect level definitions given.	OECD 408 (1991)
NOEL	Greatest concentration or amount that causes no detectable (usually adverse).....	WHO IPCS (1987)
NOEL	Highest dose… no changes distinguishable from control.	WHO IPCS (1990)
NOAEL	Highest dose… at which no toxic effects are observed.	WHO IPCS (1990)
NOEL	Highest concentration or amount… Causes no alteration (parameters and conditions defined). Replaces earlier versions?	WHO IPCS (1994)
NOAEL	Highest concentration or amount… Causes no adverse alteration (parameters and conditions defined). Replaces earlier versions?	WHO IPCS (1994)
No adverse effect level	Maximum dose or exposure level used in a test which produces no detectable signs of toxicity.	EC (1992)
NOAEL	No definition given but factors to be considered are given: Severity, time/dose response/effect, biological relevance, reversibility, normal variation and historical control.	WHO IPCS (1990, 1994)
Adverse effect	Adverse effects may be manifested as changes… which result in impairment of functional capacity or impairment of capacity to compensate for additional stress or increase in susceptibility to harmful effects of other environmental influences.	EC (1996)

Table 17: Continued

Term	Narrative definition / comments	Reference
Serious damage	Changes to the following endpoints are not regarded as 'serious': clinical observations, body-weight, food / water consumption, small changes in clinical chemistry, haematology or urinalysis, changes in organ weights with no evidence of organ dysfunction, adaptive responses (liver hypertrophy, enzyme induction, macrophage infiltration, hyperplasia in response to irritants) and where a species specific mechanism of toxicity has been demonstrated.	EC (1993)
NOAEL	An exposure level at which there are no statistically or biologically significant increases in the frequency or severity of adverse effects between the exposed population and its appropriate control. Some effects may be produced at this level, but they are not considered as adverse or precursors to adverse effects.	EPA (1995a)
Adverse effect	A biochemical change, functional impairment, or pathological lesion that either singly or in combination adversely affects the performance of the whole organism or reduces the organism's ability to respond to an additional environmental challenge.	EPA (1995a)
Biologically significant	A response in an organism or other biological system that is considered to have substantial or noteworthy effect (positive or negative) on the well being of the biological system. Used to distinguish statistically significant effects or changes, which may or may not be meaningful to the general state of health of the system.	EPA (1995a) effect
NOEL / NOAEL	Agency scientists determine the most sensitive treatment-related toxic endpoint (adverse effect) from the data submitted… This endpoint is the adverse or toxic effect that occurs in test animals at the lowest exposure to the test substance. The highest exposure that does not produce this adverse effect is called the no-observed-effect-level (NOEL) or the no-observed-adverse-effect-level (NOAEL).	FDA (1993)

Published references to terms used to describe and differentiate between adverse and non-adverse effects.

OECD

NOAEL is the abbreviation for no-observed-adverse-effect level and is the highest dose level where no adverse treatment related findings are observed.

Evident toxicity is a general term describing clear signs of toxicity following administration of test substance. These should be sufficient for hazard assessment and should be such that an increase in the dose administered can be expected to result in the development of severe toxic signs and probable mortality (OECD, 1995a).

In earlier guidelines the definition was given as: "No-effect level / no-toxic-effect level / no-adverse-effect level is the maximum dose used in a test which produces no adverse effects" (OECD, 1991, 1998).

WHO (IPCS)

"No-observed-effect level; the greatest concentration or amount of an agent, found by study or observation, that causes no detectable, usually adverse, alteration of morphology, functional capacity, growth, development, or life span of the target" (WHO IPCS, 1987).

"No-observed-effect level (NOEL): the highest dose of a substance which causes no changes distinguishable from those observed in normal (control animals)" (WHO IPCS, 1990).

"No-observed effect level: greatest concentration or amount of a substance, found by experiment or observation, that causes no alterations of morphology, functional capacity, growth, development or life span of target organisms distinguishable from those observed in normal (control) organisms of the same species and strain under the same defined conditions of exposure" (WHO IPCS, 1994).

"No-observed-adverse-effect level (NOAEL); the highest dose of a substance at which no toxic effects are observed" (WHO IPCS, 1990).

"No-observed-adverse-effect level; greatest concentration or amount of a substance, found by experiment or observation, which causes no detectable adverse alteration of morphology, functional capacity, growth, development or life span of the target organism under defined conditions of exposure alterations or morphology, functional capacity, growth, development or life span of the target may be detected which are judged not to be adverse" (WHO IPCS, 1994).

EC

"No-adverse-effect level is the maximum dose or exposure level used in a test that produces no detectable signs of toxicity" (EC, 1992).

EC guidance for risk assessment cites the following: "The no observed adverse effect level is the greatest concentration or amount of a substance found by experiment or observation, which causes no detectable adverse alteration of morphology, functional capacity, growth, development, or life span of the target organism under defined conditions of exposure.

Adverse effects may be manifested as changes in morphology, physiology, growth, development or life span of an organism which result in impairment of functional capacity or impairment of capacity to compensate for additional stress or increase in susceptibility to harmful effects of other environmental influences.''

''In the identification of the NOAEL, other factors need to be considered, such as the severity of the effect, the presence or absence of a dose- and time-effect relationship, and / or a dose- and time-response relationship, the biological relevance of an effect, the reversibility of an effect, and the normal biological variation of an effect such as may be shown by representative historical control values" (WHO IPCS, 1990; EC, 1996).

The EC's guidance on the classification of substances (EC, 1993) regarding the application of the risk phrase R48 also throws some further light on the definition of 'no adverse effect':

''Evidence indicating that R48 should not be applied. The use of this risk phrase is restricted to 'serious damage to health by prolonged exposure'. A number of substance-related effects may be observed in both humans and animals that would not justify the use of R48. These effects are relevant when attempting to determine a no-effect level for a chemical substance. Examples of well documented changes which would not normally justify classification with R48, irrespective of their statistical significance, include:

a) Clinical observations or changes in bodyweight gain, food consumption or water intake, which may have some toxicological importance but which do not, by themselves, indicate 'serious damage';
b) small changes in clinical biochemistry, haematology or urinalysis parameters which are of doubtful or minimal toxicological importance;
c) changes in organ weights with no evidence of organ dysfunction;
d) adaptive responses (e.g. macrophage migration in the lung, liver hypertrophy and enzyme induction, hyperplastic response to irritants). Local effects in the skin produced by repeated dermal application of a substance which are more appropriately classified with R38 'irritating to skin';
e) where a species-specific mechanism of toxicity (e.g. specific metabolic pathways) has been demonstrated.''

EPA

No-observed-effect-level (NOEL) is the dose level (quantity) of a substance administered to a group of experimental animals which demonstrates the absence of adverse effects observed or measured at higher dose levels.

This NOEL should produce no biologically significant differences between the group of chemically exposed animals and an unexposed control group of animals maintained under identical conditions (EPA, 1985a).

"No-effect level / no-toxic-effect level / no-adverse-effect level / no-observed-effect level" is the maximum dose used in a test, which produces no observed adverse effects (EPA, 1984).

With respect to reproduction studies, the EPA notes: "In a classic teratology study, development toxicity is restricted to adverse effects manifested in the developing organism prior to parturition. Therefore, NOELs for adverse effects which were observed in classical teratology studies (studies which encompass the embryonic and foetal periods of the organism) are more accurately referred to as 'developmental toxicity (embryo / fetotoxicity) NOELs' while adverse effects which are observed in post-natal studies should be referred to as developmental toxicity (post-natal) NOELs" (EPA, 1985b).

In EPA's Rejection Rate Analysis Report, the section relating to Guideline 82-1(a) notes that in 90-day feeding studies in rodents "The main rejection factor cited is that a NOEL was not established." Latitude should be granted by the Agency if the biological response found at the lowest dose tested in the study is either not adverse (no-observed-adverse effect level) or if NOELs can be set using longer term studies. Discussion revolves around whether a biological response is adverse (EPA, 1993).

More recent EPA definitions include:

NOAEL: "An exposure level at which there are no statistically or biological significant increases in the frequency or severity of adverse effects between the exposed population and its appropriate control. Some effects may be produced at this level, but they are not considered as adverse, nor precursors to adverse effects. In an experiment with several NOAELs, the regulatory focus is primarily on the highest one, leading to the common usage of the term NOAEL as the highest exposure without adverse effect."

Adverse effect: "A biochemical change, functional impairment, or pathological lesion that either singly or in combination adversely affects the performance of the whole organism or reduces an organism's ability to respond to an additional environmental challenge".

Biologically significant effect: "A response in an organism or other biological system that is considered to have substantial or noteworthy effect (positive or negative) on the well-being of the biological system. Used to distinguish statistically significant effects or changes, which may or may not be meaningful to the general state of health of the system" (use of Benchmark Dose approach) (EPA, 1995a).

Recent EPA Risk Assessment Guidelines cite the following:

Reproduction: "The low dose for the study is generally a NOAEL for adult and offspring effects, although if the low dose produces a biologically or statistically significant increase in response it is considered a LOAEL".

For example:- "While there is always a question as to whether weight reduction is a permanent or transitory effect, little is known about the long-term consequences of short-term foetal or neonatal weight changes. Therefore, when significant weight reduction effects are noted they are used as a basis to establish the NOAEL" (EPA, 1991).

Neurotoxicity: The NOEL is defined as "the highest dose at which there is no statistically or biological significant increase in the frequency of an adverse neurotoxic effect when compared with the appropriate control group in a database characterised as having sufficient evidence for use in a risk assessment" (EPA, 1995b).

TSCA Testing Guidelines:

"No-effect level / no-toxic-effect level / no-adverse-effect level / no-observed-effect level is the maximum dose used in a test, which produces no observed adverse effects" (EPA, 1985c).

Congress of the US Office of Technology assessment:

"The NOEL is that dose at or below which no biological effects of any type are noted (a determination that is influenced by the sensitivity of analytical techniques), and the NOAEL is that dose at or below which no harmful effects are seen. Definitions of 'harmful' effects are influenced by social norms and values. If more than one effect is seen in animal tests, the effect occurring at the lowest dose in the most sensitive animal species and sex is generally used as the basis for estimating a NOEL or NOAEL".

FDA Food Additive Testing Guidelines:

No-observed-effect level (NOEL)

"Non-treatment-related variations in the incidence of toxic endpoints occur and may depend on as number of factors, including the source of the animals, sex, genetic variations, diet, age at death, environmental conditions and the histological criteria used by pathologists.

However, Agency scientists determine the most sensitive treatment-related toxic endpoint (adverse effect) from the data submitted in support of the petition. This endpoint is the adverse or toxic effect that occurs in test animals at the lowest exposure to the test substance. The highest exposure that does not produce this adverse effect is called the no-observed-effect level (NOEL) or the no-observed-adverse-effect level (NOAEL) (FDA, 1993).

JAPAN

Japanese testing guidelines for pesticides and industrial chemicals mention the requirement for the determination of a no observable effect level, but give no definition (SACI, 1985; MITI, 1992).

Definition of NOEL (Japanese interpretation for pharmaceuticals): "Maximum dose level without toxicological adverse effects".

Criteria related to 'non-adverse effects':

1. Without histopathological changes;

2. without irreversible changes;

3. even with minor reversible pharmacological effects

This statement regarding pharmaceuticals is supported by the available guidelines, which state: "The dose level at which no toxic changes are observed - no-observed-effect level. It is noted that it is important to ascertain whether the reaction is a change anticipated from the pharmacological properties of the test substance or an unexpected response, and whether it is a reversible change or an irreversible one" (MHW, 1990).

In some cases specific test guidelines or other documents from expert groups can offer some advice on definitions:

OECD Test Guidelines 416: Two-generation reproduction toxicity study (OECD,1996)

No specific definition of adverse / non-adverse but some guidance given on what aspects to consider in interpretation of results. "The findings…should be evaluated in terms of the observed effects. The evaluation will include the relationship, or lack thereof, between the dose of the test substance and the presence or absence, incidence and severity of abnormalities, including gross lesions, identified target organs, affected fertility, clinical abnormalities, affected reproductive and litter performance, bodyweight changes, effects on mortality and any other toxic effects".

It is interesting to note that no quantitative limits or guidelines are quoted on the magnitude of any change or altered incidence.

EPA Reproductive Toxicity Risk Assessment Guidelines (EPA, 1996)

"No definition of a NOEL / NOAEL is given in this document; however there is a wealth of detail on endpoints relevant to reproductive toxicity and adverse changes".

Unfortunately in addressing the endpoints the authors have not given any guidance on the magnitude of change that would be considered adverse.

General comment in document: "Because of the limitations associated with the use of the NOAEL, the Agency is beginning to use an additional approach, the benchmark dose approach, for a more quantitative dose-response evaluation when allowed by the data".

"Although statistical analyses are important in determining the effects of a particular agent, the biological significance of data is most important. It is important to be aware that when many endpoints are investigated, statistically significant differences may occur by chance. On the other hand, apparent trends with dose may be biologically relevant even though pair-wise comparisons do not indicate a statistically significant effect".

Comments on specific endpoints:

Male reproductive organs. "Significant changes in absolute or relative organ weights may constitute an adverse reproductive effect. Significant and biologically meaningful histopathologic damage in excess of the level seen in control tissue of any of the male reproductive organs should be considered an adverse reproductive effect. Although thorough histopatholigic evaluations that fail to reveal any treatment related-effects may be quite convincing, consideration should be given to the possible presence of other testicular or epididymal effects that are not detected histologically but may affect reproductive function".

Sperm parameters. Where there is a lack of scientific insight or consensus regarding changes to endpoints, the Agency adopts a 'safety first' approach. "…the conservative approach should be taken that, within the limits indicated in the sections on those parameters, statistically significant changes in measures of sperm count, morphology, or motility as well as the number of normal sperm should be considered adverse effects".

Female reproductive organs / cyclicity and senescence. "Effects on the uterus that may be considered adverse include significant dose-related alteration or histologic abnormalities …significant increases in the rate of follicular atresia, evidence of oocyte toxicity, interference with ovulation, or altered corpus luteum formation of function should be considered adverse effects significant evidence that the oestrus cycle has been disrupted should be considered an adverse effect...significant effects on measures showing a decrease in the age of onset of reproductive senescence in females should be considered adverse".

ECETOC. Monograph on practical concepts for dose selection in chronic toxicity and carcinogenicity studies in rodents (ECETOC, 1996).

No specific definition of adverse / non-adverse but guidance is given on what aspects to consider in interpretation of results. A list of endpoints (criteria) is provided, changes to which are considered important in defining 'toxicity' (i.e. adverse effects), at the highest dose tested. "Mortality, clinical signs, reduction in bodyweight gain, haematological parameters, clinical chemistry, physiological function, and cytotoxicity".

No quantitative limits or guidelines are quoted on the magnitude of any change or altered incidence.

APPENDIX 2

Excerpt from Principles and Methods for Acute and Subchronic Toxicity (Chang *et al*, 1982)

''No Observed Effect Level (NOEL)

Since one of the main objectives of conducting subchronic studies is to define the so-called 'no effect level, this term needs to be elaborated upon. Assuming that there is a dose-effect relationship regarding some parameter(s), there must then be a dose level low enough that no adverse effects can be detected. Or put another way, the no effect level is the maximum dose that the animal can tolerate over a specific period of time without showing any adverse effects, and above which adverse effects(s) are manifested. If this is the definition of a no effect level, the question to be asked next would be which biological effects are adverse and which are not.

An adverse effect in its simplest meaning implies an abnormal, undesirable, or harmful effect to the animal's life or well-being that may be indicated by some measurable endpoint such as mortality, food consumption, body and organ weights, enzyme levels, or pathologic findings. Whether certain changes are significantly deviated from normality is often defined statistically at acceptable error levels such as $p \leq 0.05$. This value means that if the study were conducted 100 times, there are less than five times that a real no difference is wrongly concluded, on the basis of experimental data, as a significant difference (the type I error). Therefore the p value is no more than a criterion for declaring significance. It is actually the probability of making type I errors. There are also times when one may conclude that a real difference is *not* significantly different (the type II error). The smaller the p value is, the less frequently type I errors will occur, but the more frequently type II errors will occur.

In general, there are two types of significant biologic responses. First, there are the normal biologic responses, which will manifest in response to stress e.g. sweating in exercise, losing weight when starving. These changes often represent normal homeostatic reactions to stimuli. Second, there are the abnormal biologic responses, which may be caused by chemicals or other stresses, e.g. blood dryscrasia, hepatotoxicity, renal toxicity, tumours etc. Either of these biologic responses could be significantly different from the normal baseline when analysed by statistics, but obviously, it is the latter that are of more concern to toxicologists. Therefore, one must be cautious in relating a statistical finding to a true adverse biological effect.

A statistically significant finding may not automatically constitute a biologically adverse effect or a toxicologically significant effect. The magnitude of departure from the normal range, the consistency of the out-of-range responses, and the relationships of the abnormal responses to the physiological, physical, biochemical and metabolic well-being of an animal all have to be considered. It should also be noted that for many biologic parameters (e.g. haematology and clinical chemistry parameters), there exists a normal range of value, and it is possible that a value from a chemically treated group of animals will be statistically significantly different from a value of a control group of animals and still be within the so-called normal range. The final judgement should rely heavily on the toxicologist's experiences and on a case-by-case basis.

Nonetheless, the following criteria may provide general guidelines to determine if an effect is truly adverse. An effect may be considered adverse if it causes functional or anatomical impairments, causes irreversible damage to the homeostasis of the animal, increases the susceptibility of the animal to other chemical or biological insults such as infectious diseases, or causes abnormal or harmful effects on enzyme systems.

Contrasted to adverse effects, non-adverse effects can be defined as those biological effects, which do not cause physical, physiological, behavioural and biochemical changes that affect the general well-being, growth, development or life span of an animal. Furthermore, a nonadverse affect will disappear when the exposure is withdrawn from the animal.

Still, there are concerns if a true no effect level exists and perhaps if a no effect level has any real meaning. There are complicated factors involved: dose related pharmacokinetics (e.g. low dose versus high dose), the nature of the response (e.g. a single versus multiple hit response), the time of the response, and the number of animals in the study (e.g. with a 95% confidence level, even when no response is observed in 1000 animals, there is still an upper limit probability that three animals per 1000 treated would show a response). The no effect level is in fact a NOEL. It does not imply that no effect occurred, but rather that no effect was observed within the limits of the study. Different mathematical models have been suggested for the low-dose extrapolation. Readers are referred to a recent review article on quantitative risk assessment by the Committee of Food Safety Council".

BIBLIOGRAPHY

Ashby J, Lefevre PA. 2000. The prepubertal male rat assay as an alternative to the Hershberger castrated male rat assay for the detection of anti-androgens, oestrogens and metabolic modulators. *J Appl Toxicol* 20: 35-47.

Burger GT, Renne RA, Sagartz JW, Ayres PH, Coggins CRE, Mosberg AT, Hayes AW. 1989. Histologic changes in the respiratory tract induced by inhalation of xenobiotics: Physiologic adaptation or toxicity? *Toxicol Appl Pharmacol* 101: 521-542.

Carmichael NG, Enzmann H, Pate I, Waechter F. 1997. The significance of mouse liver tumor formation for carcinogenic risk assessment: results and conclusions from a survey of ten years of testing by the agrochemical industry. *Environ Health Perspect* 105: 1196-1203.

Chang PK, O'Hara GP, Hayes WA. 1982. Principles and methods of acute and subchronic toxicity. In Hayes WA ed, *Principles and methods of toxicology*. Raven Press, New York, pp 1-52.

ECETOC. 1996. Practical concepts for dose selection in chronic toxicity and carcinogenicity studies in rodents. Monograph No. 25.

ECETOC. 2003a. Toxicological modes of action: Relevance for human risk assessment. In preparation.

ECETOC. 2003b. Deviration of assessment factors for human health risk assessment. Technical Report No. 86.

EC. 1992. Commission Directive 92/69/EEC of 31 July adapting to technical progress for the seventeenth time Council Directive 67/548/EEC on the approximation of laws, regulations and administrative provisions relating to the classification, packaging and labelling of dangerous substances. *O J Eur Comm* L383A (29.12.92) p 107.

EC. 1993. Commission Directive 93/21/EEC of 27 April 1993 adapting to technical progress for the eighteenth time Council Directive 67/548/EEC on the approximation of laws, regulations and administrative provisions relating to the classification, packaging and labelling of dangerous substances *O J Eur Comm* L110A (4.5.93) pp 55-56.

EC. 1996. Technical guidance document in support of Commission Directive 93/67/EEC on risk assessment for new notified substances and Commission Regulation (EC) No 1488/94 on risk assessment for existing substances. Office for Official Publications of the EC, Luxembourg.

EPA. 1984. Pesticide assessment guidelines. Subdivision F PB86-108958.

EPA. 1985a. SEP. Toxicity potential. Guidance for analysis and evaluation of subchronic and chronic exposure studies. EPA-540/9-85-020 June 1985.

EPA. 1985b. SEP: Teratology studies. PB86-129392 June 1985.

EPA. 1985c. TSCA testing guidelines. *Federal Register* 50: 39407.

EPA. 1991. Risk assessment guidelines. *Federal Register 56:* 63798-63826

EPA. 1993. Pesticide registration rejection rate analysis: Toxicology. EPA 738-R-93-004 July 1993.

EPA. 1995a. Health risk assessment. EPA / 630 / R-94 / 007 February 1995.

EPA. 1995b. SEP. Neurotoxicity testing. *Federal Register* 60: 52032-52056.

EPA. 1996. Reproductive toxicity risk assessment guidelines. *Federal Register* 61: 56274-56332.

EPA. 2000. Profenofos - Interim re-registration eligibility decision.

FDA (Food and Drug Administration). 1993. Toxicological principles for the safety assessment of direct food additives and color additives used in food ("Red Book II". Springfield, VA, USA.

Firriolo JM, Morris CF, Trimmer GW, Twitty LD, Smith JH, Freeman JJ. 1995. Comparative 90-day feeding study with low-viscosity white mineral oil in Fischer-344 and Sprague-Dawley-derived CRL:CD rats. *Toxicol Pathol* 23: 26-33.

Fleming KA, Zimmerman H, Shubik P. 1998. Granulomas in the livers of humans and Fischer rats associated with the ingestion of mineral hydrocarbons: a comparison. *Reg Toxicol Pharmacol* 27: 75-81.

Greim H, Borm P, Schins R, Donaldson K, Driscoll K, Hartwig A, Kuempe E, Oberdorster G, and Speit G. 2001. Toxicity of fibers and particles. Report of the workshop held in Munich, Germany, 26-27 October 2000. *Inhalation Toxicology* 13: 737-754.

Hamada C, Yoshino K, Abe I, Matsumoto K, Nomura M, Yoshimura I. 1998. Detection of an outlier and evaluation of its influence in chronic toxicity studies. *Drug Info* J 32:201-212.

Haseman JK, Huff J, Boorman GA. 1984 Use of historical control data in carcinogenicity studies in rodent. *Toxicol Pathol* 12: 126-135.

MHW (Japanese Ministry for Health and Welfare). 1990. Guidelines for toxicity studies of drugs.

Lang PL. 1992. Spontaneous neoplastic lesions and selected non-neoplastic lesions in the Crl:CD® BR Rat. Charles River Laboratories. Charles River publication.

Miller J. 1995. Uptake and fate of ozone in the respiratory tract. *Toxicol Lett* 82 / 83: 277-285.

MITI. 1992. Testing Guidelines.

Morgan KT. 1997. A brief review of formaldehyde carcinogenesis in relation to rat nasal pathology and human health risk assessment. *Toxicol Pathol* 25: 291-307.

OECD. 1991 (updated 1998). Guideline for Testing of Chemicals. 408. Repeated dose 90-day oral toxicity study in rodents.

OECD. 1995a. Guideline for Testing of Chemicals. 407. Repeated dose 28-day oral toxicity study in rodents.

OECD. 1995b. Guideline for Testing of Chemicals. 421.Reproduction / developmental toxicity screening test.

OECD. 1996. Guidelines for the Testing of Chemicals (draft) (updated 2001). 416. Two generation reproduction toxicity study.

Paynter OE. 1984. Oncogenic potential: guidance for analysis and evaluation of long term rodent studies. Evaluation procedure #1000.1 Office of Pesticide and Toxic Substances, EPA, Washington, DC.

Quest JA. 1990. Pesticide residues in food. Joint FAO / WHO meeting on pesticide residues. Toxicology evaluations 1990. ICPS / International Program on Chemical Safety. Profenofos, pp. 123-150.

SACI (Society of Agricultural Chemical Industry). 1985. Agricultural Chemicals Laws and Regulations Japan (II).

Shoda T, Toyoda K, Uneyama C, Takada K, Takahashi M. 1997. Lack of carcinogenicity of medium-viscosity liquid paraffin given in the diet to F344 rats. *Fd Chem Toxic* 35: 1181-1190.

Tyl RW, Myers CV, Marr MC, Brine DR, Fail PA, Seely JC, Van Miller JP. 1999. Two-generation reproductive study with para-octylphenol in rats. *Regul Toxicol Pharmacol* 30: 81-95.

Ward JM, Uno H, Frith CH. 1993. Immunohistochemistry and morphology of reactive lesions in lymph nodes and spleen from rats and mice. *Toxicol Pathol* 21: 199-205.

WHO IPCS (International Program of Chemical Safety). 1987. Environmental Health Criteria No 70. Principles for the safety assessment of food additives and contaminants in food. World Health Organization, Geneva, Switzerland.

WHO IPCS (International Program of Chemical Safety). 1990. Environmental Health Criteria No 104. Principles for the toxicological assessment of pesticide residues in food. World Health Organization, Geneva, Switzerland.

WHO IPCS (International Program of Chemical Safety). 1994. Environmental Health Criteria Assessing Human Health Risks of Chemicals No 170. Derivation of guidance values for health-based exposure limits. World Health Organization, Geneva, Switzerland.

WHO PCS (Program for Promotion of Chemical Safety). 1998. Report of a consultation on interpretation of inhibition of acetylcholinesterase activity (PCS / 98.7).

Willard MD, Twedt DC. 1994. Gastrointestinal, pancreatic, and hepatic disorders. In Willard MD, Tvedten H, Turntable GH, eds,*Small animal clinical diagnosis by laboratory methods*. 2nd ed.

MEMBERS OF THE TASK FORCE

R. Lewis (Chairman)	Syngenta UK – Macclesfield
R. Billington	Dow AgroScience UK – Wantage
E. Debruyne	Bayer CropScience F – Sophia Antipolis
A. Gamer	BASF D – Ludwigshafen
B. Lang	Syngenta CH – Basel
J. Van Miller*	Union Carbide USA – Danbury
F. Carpanini (Secretary)	ECETOC

* Now at Toxicology / Regulatory Services Inc., USA - Charlottesville

MEMBERS OF THE SCIENTIFIC COMMITTEE

B. Hildebrand (Chairman)	Consultant D – Weinheim
N. Carmichael (Vice-chairman) Head, Toxicology	Bayer CropScience F – Sophia Antipolis
G. Randall (Vice-chairman) Director, Environmental Laboratory	AstraZeneca UK – Brixham
C. d'Hondt Head, Environmental Safety Department	Syngenta CH – Basel
E. Bomhard Industrial Toxicology	Bayer D – Wuppertal
C. Braun Occupational Toxicologist	Akzo Nobel NL – Arnhem
P. Calow Professor of Zoology	University of Sheffield UK – Sheffield
P. Douben Head, SEAC Environmental Protection Department	Unilever UK – Bebington
T. Feijtel Manager, Professional and Regulatory Services	Procter & Gamble B – Brussels
H. Greim Director, Institute of Toxicology and Environmental Hygiene	Technical University Munich D – Munich
J. Jackson Senior Associate, Medical Adviser	Monsanto B – Brussels
C. Money Industrial Hygiene Adviser – Europe	ExxonMobil B – Brussels
A. Sarrif Director, Toxicology Affairs, Europe	DuPont B – Brussels
G. Swaen Head, Occupational Epidemiology Unit	Maastricht University NL – Maastricht
B. van Ravenzwaay Director, Experimental Toxicology and Ecology	BASF D – Ludwigshafen
H-J. Wiegand Head, Product Safety Department	Degussa D – Düsseldorf

Stewards responsible for primary peer review:
N. Carmichael
W. Tordoir
B. Hildebrand

ECETOC PUBLISHED REPORTS

Monographs

No.	Title
No. 1	Good Laboratory Practice
No. 2	A Contribution to Strategy for Identification and Control of Occupational Carcinogens
No. 3	Risk Assessment of Occupational Chemical Carcinogens
No. 4	Hepatocarcinogenesis in Laboratory Rodents: Relevance for Man
No. 5	Identification and Assessment of the Effects of Chemicals on Reproduction and Development (Reproductive Toxicology)
No. 6	Acute Toxicity Tests, LD50 (LC50) Determinations and Alternatives
No. 7	Recommendations for the Harmonisation of International Guidelines for Toxicity Studies
No. 8	Structure-Activity Relationships in Toxicology and Ecotoxicology: An Assessment (Summary)
No. 9	Assessment of Mutagenicity of Industrial and Plant Protection Chemicals
No. 10	Identification of Immunotoxic Effects of Chemicals and Assessment of their Relevance to Man
No. 11	Eye Irritation Testing
No. 12	Alternative Approaches for the Assessment of Reproductive Toxicity (with emphasis on embryotoxicity / teratogenicity)
No. 13	DNA and Protein Adducts: Evaluation of their Use in Exposure Monitoring and Risk Assessment
No. 14	Skin Sensitisation Testing
No. 15	Skin Irritation
No. 16	Early Indicators of Non-Genotoxic Carcinogenesis
No. 17	Hepatic Peroxisome Proliferation
No. 18	Evaluation of the Neurotoxic Potential of Chemicals
No. 19	Respiratory Allergy
No. 20	Percutaneous Absorption
No. 21	Immunotoxicity: Hazard Identification and Risk Characterisation
No. 22	Evaluation of Chemicals for Oculotoxicity
No. 23	Receptor Mediated Mechanisms in Chemical Carcinogenesis
No. 24	Risk Assessment for Carcinogens
No. 25	Practical Concepts for Dose Selection in Chronic Toxicity and Carcinogenicity Studies in Rodents
No. 26	Aquatic Toxicity Testing of Sparingly Soluble Volatile and Unstable Substances
No. 27	Aneuploidy
No. 28	Threshold-Mediated Mutagens - Mutation Research Special Issue
No. 29	Skin Sensitisation Testing for the Purpose of Hazard Identification and Risk Assessment
No. 30	Genetic Susceptibility to Environmental Toxicants
No. 31	Guidance on Evaluation of Reproductive Toxicity Data
No. 32	Use of Human Data in Hazard Classification for Irritation and Sensitisation

Technical Reports

No.	Title
No. 1	Assessment of Data on the Effects of Formaldehyde on Humans
No. 2	The Mutagenic and Carcinogenic Potential of Formaldehyde
No. 3	Assessment of Test Methods for Photodegradation of Chemicals in the Environment
No. 4	The Toxicology of Ethylene Glycol Monoalkyl Ethers and its Relevance to Man
No. 5	Toxicity of Ethylene Oxide and its Relevance to Man
No. 6	Formaldehyde Toxicology: An Up-Dating of ECETOC Technical Reports 1 and 2
No. 7	Experimental Assessment of the Phototransformation of Chemicals in the Atmosphere
No. 8	Biodegradation Testing: An Assessment of the Present Status
No. 9	Assessment of Reverse-Phase Chromatographic Methods for Determining Partition Coefficients
No. 10	Considerations Regarding the Extrapolation of Biological Data in Deriving Occupational Exposure Limits
No. 11	Ethylene Oxide Toxicology and its Relevance to Man: An Up-Dating of ECETOC Technical Report No. 5
No. 12	The Phototransformation of Chemicals in Water: Results of a Ring-Test
No. 13	The EEC 6th Amendment: A Guide to Risk Evaluation for Effects on the Environment
No. 14	The EEC 6th Amendment: A Guide to Risk Evaluation for Effects on Human Health
No. 15	The Use of Physical-Chemical Properties in the 6th Amendment and their Required Precision, Accuracy and Limiting Values
No. 16	A Review of Recent Literature on the Toxicology of Benzene
No. 17	The Toxicology of Glycol Ethers and its Relevance to Man: An Up-Dating of ECETOC Technical Report No. 4
No. 18	Harmonisation of Ready Biodegradability Tests
No. 19	An Assessment of Occurrence and Effects of Dialkyl-o-Phthalates in the Environment
No. 20	Biodegradation Tests for Poorly-Soluble Compounds
No. 21	Guide to the Classification of Carcinogens, Mutagens, and Teratogens under the 6th Amendment
No. 22	Classification of Dangerous Substances and Pesticides in the EEC Directives. A Proposed Revision of Criteria for Inhalational Toxicity
No. 23	Evaluation of the Toxicity of Substances to be Assessed for Biodegradability
No. 24	The EEC 6th Amendment: Prolonged Fish Toxicity Tests
No. 25	Evaluation of Fish Tainting
No. 26	The Assessment of Carcinogenic Hazard for Human Beings exposed to Methylene Chloride
No. 27	Nitrate and Drinking Water
No. 28	Evaluation of Anaerobic Biodegradation
No. 29	Concentrations of Industrial Organic Chemicals Measured in the Environment: The Influence of Physico-Chemical Properties, Tonnage and Use Patterns
No. 30	Existing Chemicals: Literature Reviews and Evaluations (Fifth Edition) (No longer available)
No. 31	The Mutagenicity and Carcinogenicity of Vinyl Chloride: A Historical Review and Assessment
No. 32	Methylene Chloride (Dichloromethane): Human Risk Assessment Using Experimental Animal Data
No. 33	Nickel and Nickel Compounds: Review of Toxicology and Epidemiology with Special Reference to Carcinogenesis
No. 34	Methylene Chloride (Dichloromethane): An Overview of Experimental Work Investigating Species Differences in Carcinogenicity and their Relevance to Man
No. 35	Fate, Behaviour and Toxicity of Organic Chemicals Associated with Sediments
No. 36	Biomonitoring of Industrial Effluents
No. 37	Tetrachlorethylene: Assessment of Human Carcinogenic Hazard
No. 38	A Guide to the Classification of Preparations Containing Carcinogens, Mutagens and Teratogens

No. 39 Hazard Assessment of Floating Chemicals After an Accidental Spill at Sea
No. 40 Hazard Assessment of Chemical Contaminants in Soil
No. 41 Human Exposure to N-Nitrosamines, their Effects and a Risk Assessment for N-Nitrosodiethanolamine in Personal Care Products
No. 42 Critical Evaluation of Methods for the Determination of N-Nitrosamines in Personal Care and Household Products
No. 43 Emergency Exposure Indices for Industrial Chemicals
No. 44 Biodegradation Kinetics
No. 45 Nickel, Cobalt and Chromium in Consumoducts: Allergic Contact Dermatitis
No. 46 EC 7th Amendment: Role of Mammalian Toxicokinetic and Metabolic Studies in the Toxicological Assessment of Industrial Chemicals
No. 47 EC 7th Amendment "Toxic to Reproduction": Guidance on Classification
No. 48 Eye Irritation: Reference Chemicals Data Bank (Second Edition)
No. 49 Exposure of Man to Dioxins: A Perspective on Industrial Waste Incineration
No. 50 Estimating Environmental Concentrations of Chemicals using Fate and Exposure Models
No. 51 Environmental Hazard Assessment of Substances
No. 52 Styrene Toxicology Investigation on the Potential for Carcinogenicity
No. 53 DHTDMAC: Aquatic and Terrestrial Hazard Assessment (CAS No. 61789-80-8)
No. 54 Assessment of the Biodegradation of Chemicals in the Marine Environment
No. 55 Pulmonary Toxicity of Polyalkylene Glycols
No. 56 Aquatic Toxicity Data Evaluation
No. 57 Polypropylene Production and Colorectal Cancer
No. 58 Assessment of Non-Occupational Exposure to Chemicals
No. 59 Testing for Worker Protection
No. 60 Trichloroethylene: Assessment of Human Carcinogenic Hazard
No. 61 Environmental Exposure Assessment
No. 62 Ammonia Emissions to Air in Western Europe
No. 63 Reproductive and General Toxicology of some Inorganic Borates and Risk Assessment for Human Beings
No. 64 The Toxicology of Glycol Ethers and its Relevance to Man
No. 65 Formaldehyde and Human Cancer Risks
No. 66 Skin Irritation and Corrosion: Reference Chemicals Data Bank
No. 67 The Role of Bioaccumulation in Environmental Risk Assessment: The Aquatic Environment and Related Food Webs
No. 68 Assessment Factors in Human Health Risk Assessment
No. 69 Toxicology of Man-Made Organic Fibres
No. 70 Chronic Neurotoxicity of Solvents
No. 71 Inventory of Critical Reviews on Chemicals (Only available to ECETOC members)
No. 72 Methyl tert-Butyl Ether (MTBE) Health Risk Characterisation
No. 73 The Value of Aquatic Model Ecosystem Studies in Ecotoxicology
No. 74 QSARs in the Assessment of the Environmental Fate and Effects of Chemicals
No. 75 Organophosphorus Pesticides and Long-term Effects on the Nervous System
No. 76 Monitoring and Modelling of Industrial Organic Chemicals, with Particular Reference to Aquatic Risk Assessment
No. 77 Skin and Respiratory Sensitisers: Reference Chemicals Data Bank
No. 78 Skin Sensitisation Testing: Methodological Considerations
No. 79 Exposure Factors Sourcebook for European Populations (with Focus on UK Data)
No. 80 Aquatic Toxicity of Mixtures
No. 81 Human Acute Intoxication from Monochloroacetic Acid: Proposals for Therapy
No. 82 Risk Assessment in Marine Environments
No. 83 The Use of T25 Estimates and Alternative Methods in the Regulatory Risk Assessment of Non-threshold Carcinogens in the European Union
No. 84 Scientific Principles for Soil Hazard Assessment of Substances

Joint Assessment of Commodity Chemicals (JACC) Reports

No.	Title
No. 1	Melamine
No. 2	1,4-Dioxane
No. 3	Methyl Ethyl Ketone
No. 4	Methylene Chloride
No. 5	Vinylidene Chloride
No. 6	Xylenes
No. 7	Ethylbenzene
No. 8	Methyl Isobutyl Ketone
No. 9	Chlorodifluoromethane
No. 10	Isophorone
No. 11	1,2-Dichloro-1,1-Difluoroethane (HFA-132b)
No. 12	1-Chloro-1,2,2,2-Tetrafluoroethane (HFA-124)
No. 13	1,1-Dichloro-2,2,2-Trifluoroethane (HFA-123)
No. 14	1-Chloro-2,2,2-Trifluoromethane (HFA-133a)
No. 15	1-Fluoro 1,1-Dichloroethane (HFA-141B)
No. 16	Dichlorofluoromethane (HCFC-21)
No. 17	1-Chloro-1,1-Difluoroethane (HFA-142b)
No. 18	Vinyl Acetate
No. 19	Dicyclopentadiene (CAS: 77-73-6)
No. 20	Tris-/Bis-/Mono-(2 ethylhexyl) Phosphate
No. 21	Tris-(2-Butoxyethyl)-Phosphate (CAS:78-51-3)
No. 22	Hydrogen Peroxide (CAS: 7722-84-1)
No. 23	Polycarboxylate Polymers as Used in Detergents
No. 24	Pentafluoroethane (HFC-125) (CAS: 354-33-6)
No. 25	1-Chloro-1,2,2,2-tetrafluoroethane (HCFC 124) (CAS No. 2837-89-0)
No. 26	Linear Polydimethylsiloxanes (CAS No. 63148-62-9)
No. 27	n-Butyl Acrylate (CAS No. 141-32-2)
No. 28	Ethyl Acrylate (CAS No. 140-88-5)
No. 29	1,1-Dichloro-1-Fluoroethane (HCFC-141b) (CAS No. 1717-00-6)
No. 30	Methyl Methacrylate (CAS No. 80-62-6)
No. 31	1,1,1,2-Tetrafluoroethane (HFC-134a) (CAS No. 811-97-2)
No. 32	Difluoromethane (HFC-32) (CAS No. 75-10-5)
No. 33	1,1-Dichloro-2,2,2-Trifluoroethane (HCFC-123) (CAS No. 306-83-2)
No. 34	Acrylic Acid (CAS No. 79-10-7)
No. 35	Methacrylic Acid (CAS No. 79-41-4)
No. 36	n-Butyl Methacrylate; Isobutyl Methacrylate (CAS No. 97-88-1) (CAS No. 97-86-9)
No. 37	Methyl Acrylate (CAS No. 96-33-3)
No. 38	Monochloroacetic Acid (CAS No. 79-11-8) and its Sodium Salt (CAS No. 3926-62-3)
No. 39	Tetrachloroethylene (CAS No. 127-18-4)
No. 40	Peracetic Acid (CAS No. 79-21-0) and its Equilibrium Solutions

Special Reports

No.	Title
No. 8	HAZCHEM; A Mathematical Model for Use in Risk Assessment of Substances
No. 9	Styrene Criteria Document
No. 10	Hydrogen Peroxide OEL Criteria Document (CAS No. 7722-84-1)
No. 11	Ecotoxicology of some Inorganic Borates
No. 12	1,3-Butadiene OEL Criteria Document (Second Edition) (CAS No. 106-99-0)
No. 13	Occupational Exposure Limits for Hydrocarbon Solvents
No. 14	n-Butyl Methacrylate and Isobutyl Methacrylate OEL Criteria Document
No. 15	Examination of a Proposed Skin Notation Strategy
No. 16	GREAT-ER User Manual

Documents

No.	Title
No. 32	Environmental Oestrogens: Male Reproduction and Reproductive Development
No. 33	Environmental Oestrogens: A Compendium of Test Methods
No. 34	The Challenge Posed by Endocrine-disrupting Chemicals
No. 35	Exposure Assessment in the Context of the EU Technical Guidance Documents on Risk Assessment of Substances
No. 36	Comments on OECD Draft Detailed Review Paper: Appraisal of Test Methods for Sex-Hormone Disrupting Chemicals
No. 37	EC Classification of Eye Irritancy
No. 38	Wildlife and Endocrine Disrupters: Requirements for Hazard Identification
No. 39	Screening and Testing Methods for Ecotoxicological Effects of Potential Endocrine Disrupters: Response to the EDSTAC Recommendations and a Proposed Alternative Approach
No. 40	Comments on Recommendation from Scientific Committee on Occupational Exposure Limits for 1,3-Butadiene
No. 41	Persistent Organic Pollutants (POPs) Response to UNEP / INC / CEG-I Annex 1
No. 42	Genomics, Transcript Profiling, Proteomics and Metabonomics (GTPM). An Introduction