ECETOC

Brussels, 11 February 1986

Technical Report

N° 21

A GUIDE TO THE CLASSIFICATION OF CARCINOGENS, MUTAGENS AND TERATOGENS UNDER THE SIXTH AMENDEMENT

ISSN 0773-8072

European Chemical Industry
Ecology & Toxicology Centre

Avenue Louise 250, B. 63
B - 1050 Brussels, Belgium

TABLE OF CONTENTS

We wish to acknowledge very valuable discussions with Professor B. Bridges of the University of Sussex and Professor F.H. Sobels of the University of Leiden who helped to clarify our ideas on the mutagenicity section of this report. The contents of the document are, however, wholly the responsibility of ECETOC.

A. SUMMARY

The Sixth Amendment to the European Communities' Council Directive 67/548/EEC relating to the classification, labelling and packaging of dangerous substances requires a manufacturer or importer to propose a classification of such substances according to their toxic potential. The classification, which is based on physico-chemical and toxicological properties, leads to labelling and packaging requirements.

In Annex VI of the 6th Amendment, clear guidance is given for classifying substances with regard to their LD_{50} or LC_{50}, and it is noted that the classification may, alternatively, have to be based on other effects. The 5th Adaptation of the Directive gives further guidance on classification with respect to acute effects and also provides definitions of 3 Categories of carcinogens, 3 of mutagens and 2 of teratogens. However, it gives no detailed guidance on the criteria which would lead to classification in these Categories. Unlike substances whose classification is solely defined by a numerically-expressed acute effect, the classification of carcinogenic, mutagenic or teratogenic substances has to be based on studies the results of which can rarely be expressed in numerical terms, plus other relevant information. This means that such classifications have to be derived by an expert interpretation of all relevant information on each individual substance.

In this report ECETOC has set out general guidance for the classification of new and existing carcinogenic, mutagenic and teratogenic substances as an aid to those who have responsibilities in this area. The evidence which leads to the inclusion in, or exclusion from, each Category, or to non-classification, is discussed.

B. INTRODUCTION

In 1979 the European Communities published Council Directive 79/831/EEC amending for the sixth time the Directive 67/548/EEC relating to the classification, packaging and labelling of dangerous substances, henceforth referred to as the "6th Amendment". This amendment has been incorporated into legislation by the member states and where there are differences in the text between the 6th Amendment and the national versions the English text as issued by the European Commission was used in producing this report.

In Article 2.1(a), substances are defined as "chemical elements and their compounds as they occur in the natural state or as produced by industry, including any additives required for the purpose of placing them on the market."

In the 6th Amendment there is a well-defined procedure for notifying the intention to place a new substance on the market. According to the tonnage to be marketed, information relevant to potential adverse effects on humans or environmental organisms has to be developed from a series of experimental measurements or studies as recommended in Annexes VII and VIII of the Amendment (see ECETOC, 1984). The manufacturer or importer has also to propose a classification of the substance, based on the results of the experimental measurements and studies, and according to the Categories referred to in Annex VI, section II.D. of the 6th Amendment. The detailed requirements for categorisation are set out in the 5th Adaptation of the Directive (83/467/EEC of 29.7.83), henceforth referred to as the "5th Adaptation". It is also required to classify substances already on the market according to the same Categories (see Annex I of the 6th Amendment).

In Annex VI, guidance is given for classifying substances with regard to their LD_{50} or LC_{50}. In part 1, A,(b) of this Annex it is stated that "if facts show that for the purposes of classification it is inadvisable to use the LD_{50} or LC_{50} values as a principal basis because the substances or preparations produce other effects the substances or preparations shall be classified according to the magnitude of these effects". This means that for some substances the classification has to reflect their other acute effects, or chronic effects, rather than the lethal dose. In the 5th Adaptation, 3 Categories of carcinogens, 3 of mutagens and 2 of teratogens are defined, and the procedure for notifying the classification is given (see Article 3.1.2). Unlike substances whose classification can be defined by a numerically-expressed acute effect (eg. by an LD_{50}), classification of carcinogenic, mutagenic or teratogenic substances has to be based on the results of toxicological or epidemiological studies which can rarely be expressed in numerical terms. It is emphasised that such classifications have to be derived by a critical evaluation of the results of the studies undertaken and an expert interpretation of all relevant information on each individual substance.

The tests or studies used for classification purposes should be compatible with the relevant OECD test guidelines and Good Laboratory Practice (GLP), or their adequacy should be judged on a case by case basis.

In this report, produced by a Task Force whose members are given in Section G, ECETOC has set out guidance for the classification of substances as carcinogens, mutagens and teratogens, as an aid to those who have responsibilities in this area.

C. CLASSIFICATION OF CARCINOGENS

In the 5th Adaptation, carcinogenic substances are required to be classified in three Categories :

"Category 1. Substances known to be carcinogenic to man. There is sufficient evidence to establish a causal association between human exposure to a substance and the development of cancer.

Category 2. Substances which should be regarded as if they are carcinogenic to man. There is sufficient evidence to provide a strong presumption that human exposure to a substance may result in the development of cancer, generally on the basis of :

- appropriate long-term animal studies,
- other relevant information.

Category 3. Substances which cause concern for man owing to possible carcinogenic effects but in respect of which the available information is not adequate for making a satisfactory assessment. There is some evidence from appropriate animal studies, but this is insufficient to place the substance in Category 2".

In the following sections the factors to be considered before substances are placed in one or other of the above Categories, or are not classified, are given.

1. Category 1 : Substances Known to be Carcinogenic to Man

   Evidence for the inclusion of a substance in this Category is provided by sound epidemiological studies. It must be recognised that such studies can

only suggest or demonstrate an association between an increased incidence of cancer in a human population and exposure to a particular substance. To demonstrate beyond reasonable doubt that exposure has actually caused cancer is difficult. The following observations support the existence of a causal relationship :

a) A marked increase in the incidence of one or more types of cancer in the exposed, compared to the non-exposed, population.
b) The occurrence of an unusually high incidence of a rare cancer in a population exposed to a substance, particularly where such tumours do not occur normally in geographical clusters.
c) The observation of an increase in cancer incidence in more than one study, carried out on populations from different locations or workplaces. This reinforces the evidence of a causal relationship suggested by the findings under a) and b), above.
d) An increase in cancer incidence associated with exposure to increasing amounts of a substance or with increasing periods of exposure prior to the development of the disease, .e. evidence of dose-time-response relationships. Supporting evidence is provided by the observation of a decrease in the incidence of cancer following a reduction in the exposure of populations to a substance.

Limitations in the epidemiological data can make it difficult to determine whether a causal relationship exists. For example :

i) the sample size may be too small to provide the statistical power necessary for detecting an excess incidence of cancer;
ii) information on the level and duration of exposure may be inadequate;
iii) the study period may be too short for a latency period to elapse;
iv) an excess incidence of tumours or an elevated death-rate among the exposed persons may be due to confounding factors (such as exposure to other substances, smoking habits or alcohol consumption) which were not, or could not be taken into account.

2. Category 2 : Substances Which Should be Regarded as if They Are Carcinogenic to Man

Evidence leading to a classification in this Category will come from appropriate long-term animal studies together with other relevant information. Substances in this Category fall into the class of "Putative

Human Chemical Carcinogens", defined by ECETOC (1982) as "a clearly-defined chemical substance which causes cancers in adequate animal experimentation, under exposure conditions which corrrespond to those in man or where the relevance of exposure conditions can be deduced". The phrase in the 5th Adaptation definition "...a strong presumption that human exposure to a substance may result in the development of cancer...", and the ECETOC definition, can be interpreted to exclude from Category 2 those substances that produce cancers only under conditions which are irrelevant to the likely human exposure, or where the relevance cannot be deduced. (see C.2.2.2).

When the adequacy of animal data for a Category 2 classification is considered, two aspects have to be taken into account : the reliability of the animal data, and the criteria for a Category 2 classification based on them. These are discussed below.

2.1 Assessment of the reliability of animal data

Generally speaking, current studies for investigating carcinogenic potential follow recognised guidelines such as those of the OECD, EPA or FDA. However, none of these guidelines can be regarded as comprising rigid requirements and the reliability of data from any study (including those carried out before these guidelines were published) should be determined on a case by case basis. The following criteria should be borne in mind when reliability is assessed :

a) The test substance should have been clearly specified and the amount administered should have been monitored. For example, during inhalation studies the concentration of the test substance and the characteristics of the atmosphere (e.g. air flow, concentration, temperature, humidity and, where relevant, particle size) should have been carefully monitored, and the substance should have been administered in a form similar to that involved in human exposure. If the test substance was administered in the feed, its concentration, homogeneity and stability in the feed should have been determined.

b) The number of animals (normally rodents) used should have been high enough to identify a small increase in neoplasms above the background incidence (50 per sex, per dose in the exposed and control groups are usually required).

c) Ideally both sexes should have been examined. This may not be necessary in some types of study, e.g. in skin painting in mice to assess topical carcinogenicity, for which the use of a single sex is generally regarded as acceptable because there are substantial background data on such studies.

d) When there has been no increase in the incidence of tumours, the highest dose level should have produced some toxic effects without significantly reducing the survival. Where there is an increase in tumour incidence, the survival of test animals should not have been significantly less than that of the controls, unless the animals died early because of treatment-related tumours.

e) The control groups should have been of the same size as, or larger than, the dosed groups, and treated in the same way as the dosed groups except for the absence of test material. In specific cases it may be necessary to include a positive-control group to demonstrate the susceptibility of the animal strain used.

f) The substance should have been administered, either continuously or with adequate frequency, over a high proportion of the total life span of the animal.

g) Because of the variability of tumour incidence in control animals, adequate background data should have been available to assist in interpreting the results.

Some substances are sufficiently potent to provide adequate evidence of carcinogenicity even if many of the above criteria are not fulfilled, eg. the use of fewer animals is adequate when there is a clear increase in the incidence of tumours in the exposed as compared to control animals.

2.2 Criteria for a Category 2 classification based on findings from animal studies.

2.2.1 Inclusion criteria. For a classification in Category 2 it is recommended that at least one of the following conditions described by the International Agency for Research on Cancer (IARC, 1984) should be met, ie. there should be "an increased incidence of malignant tumours : (a)

in multiple species or strains, or (b) in multiple experiments (preferably with different routes of administration or using different dose levels); or (c) to an unusual degree with regard to incidence, site or type of tumour, or age at onset. Additional evidence may be provided by data on dose-response effects".

The criteria for including a substance in this Category are based upon finding an increased incidence of malignant tumours (c.f. 2.2.1.). There are a few instances, however, where a precise distinction between benign and malignant tumours is not possible, eg. for certain endocrine or skin tumours. In this case, providing also that a progression from the apparently benign state to a more overtly malignant one is known for such tumours, it may be valid to combine them when assessing tumour incidence in a particular organ. Furthermore, where benign and malignant tumours occur together in a particular study and originate from one and the same cell type in an organ or tissue, it may be valid to combine them in assessing tumour incidence. Nevertheless, the finding of an increase in malignant tumours is the hallmark of carcinogenicity and for this reason it is stressed in the text.

2.2.2 Exclusion criteria. Data from an animal study would not normally be regarded as sufficient for a Category 2 classification if :

a) the route of exposure or physical form of the substance was not relevant to the human situation, or its relevance could not be deduced;
b) the dose level required to produce tumours in experimental animals was so high that it adversely affected their physiological state or produced recurrent injury in the tissue or organ in which tumours later appeared, but no tumours developed at exposure levels at which these effects were not produced ;
c) tumours were produced in animals following repeated administration of a substance only at levels which would require the intake of unrealistically large quantities by man or under conditions which could not be tolerated by man;
d) only benign tumours were induced and there was no evidence of progression to malignancy;
e) there was an increased incidence only of tumours with a high spontaneous incidence in the species used, eg. liver, lung, lymphatic and mammary tumours in the mouse ;
f) only hepatic nodules were induced in rats;

g) other relevant information raised doubts about the significance of the animal data to man.

In such cases the substance should be considered for classification in Category 3.

2.3 The use of other relevant information

In this context the phrase "Other relevant information" in the 5th Adaptation is interpreted to mean information other than that from the long-term animal studies. For the purpose of classification such information should not normally be used in place of data from animal studies but rather as an aid in interpreting them, especially if it helps in deciding whether humans are likely to be affected.

a) Epidemiological data.

i) The results of epidemiological studies which are inadequate to support a Category 1 classification might support a classification in Category 2, but only when the animal data are largely sufficient on their own.

ii) Adequate epidemiological studies which provide good evidence for the absence of carcinogenicity should be taken into account when the relevance of data from animal studies to man is assessed.

b) Results from short-term tests.

These tests should be such as to enable the ability of a substance to interact with DNA to be assessed. The interaction of a chemical with DNA, which may result in gene, chromosome or genome mutations in somatic cells, is thought to be a necessary step in the initial stages of chemical carcinogenesis. Many test systems have been developed for assessing the mutagenic (genotoxic) activity of a substance but the extent to which these tests have been validated as indices of carcinogenic potential varies. The results from an appropriate battery of adequately validated short-term tests may assist classification.

Substances which are clearly mutagenic and possess carcinogenic activity according to the criteria outlined above will obviously be placed in Category 2. Substances which, although positive in animal carcinogenicity studies, give negative results in an appropriate battery of short-term tests covering the above end-points may not appropriately be classified in Category 2 if the exposure levels at which tumours are

induced in the animal studies greatly exceed the likely level of human exposure.

The demonstration of DNA-adduct formation or chromosome damage in exposed human populations does not prove that the substance is a human carcinogen, but in conjunction with positive results from animal studies it could support a Category 2 classification.

c) Data concerning the differences and similarities in metabolism, absorption and toxicokinetics between man and test species may be useful in assessing the relevance to man of the results of animal studies.

d) Structure-activity relationships (SARs). These are not at present reliable enough to support or exclude the classification of a substance as a carcinogen, although in a few cases where particularly strong SARs have been established for series of structurally-related chemicals. eg. the alkyl nitrosamines, they may suggest that a substance possibly has carcinogenic potential.

3. Category 3 : Substances Which Cause Concern Owing to Possible Carcinogenic Effects

Substances for which the limited evidence of carcinogenicity does not meet the criteria for classification in Categories 1 or 2 as described above should be considered for this Category.

A substance should be classified in Category 3 only after an expert consideration of all the available evidence. The following circumstances, taken singly or together, exemplify such evidence :

a) there is an increase in the incidence of malignant tumours but it is only of borderline significance ;
b) there is a treatment-related increase only of benign tumours;
c) there is an increased incidence of malignant tumours in only 1 species or strain but the criteria for classification in Category 2 (C.2.2.1) are not met;
d) there is an increased incidence of tumours only in organs or tissues of which the natural incidence is high, or known to be variable;
e) there is an increased incidence of malignant tumours induced only by a route of administration which is not relevant to human exposure, or where the relevance cannot be deduced. It should be noted that some

routes of administration (e.g. subcutaneous, intravenous and intraperitoneal) may be less reliable indicators of carcinogenicity than are others. The subcutaneous route, in particular, can give false positive findings and much caution must be exercised in interpreting the data, especially when only local tumour formation is induced;

f) there is a clear increase in the incidence of malignant tumours but only at exposure levels which are not relevant to the human situation, or where the relevance cannot be deduced;

g) other information suggests that positive findings in animal studies are not relevant to humans, for example when extensive epidemiological studies provide a strong indication that the carcinogenic risk to man is negligible;

h) adequately-performed animal studies at exposure levels relevant to that of humans have given results which do not meet the criteria for classification in Category 2, and genotoxic activity has not been demonstrated in an appropriate battery of short-term tests (see section C.2.3.b).

4. Circumstances Leading to Non-classification

In the absence of evidence for carcinogenicity, a substance is obviously not classified. There are also circumstances when the evidence is not sufficient for classification, as in the following examples.

a) The experimental animals were exposed under conditions which do not occur when the substance is handled or used by humans (eg. "solid state" sarcomas from the implantation of a plastic), or which involved the use of the substance in a physical form to which humans are not exposed.

b) There is only equivocal evidence of carcinogenicity from animal studies where, i) it is considered that the observed tumours probably did not occur as a result of exposure to the substance; or ii) where the experimental studies were either inadequately performed or could not be reproduced.

c) Tumours were produced in animals by a substance which led to :

   i) a marginal but statistically-significant increase in the incidence of tumours only at dose levels which greatly exceed those likely to be experienced by man, especially if the increase appeared towards the end of the animals' life-span;

ii) an increase in the incidence of tumours only at dose levels which produced recurrent toxic damage in the target tissue.

The criteria under c), above, are especially important if the substance lacks mutagenic activity, and considerable experience is required to assess the significance of such findings to man. These findings are exemplified by substances which cause ;

- recurrent tissue damage leading to tissue regeneration (eg. skin painting with highly irritant materials);
- an alteration in the physiological state of the animal. This may be due to the physico-chemical properties of the substance, eg. which cause its deposition as bladder stones in the rat, leading to the development of bladder cancer. Other substances administered at high dose levels may produce dietary changes or alterations in hormonal balance, both of which can influence the incidence of spontaneous tumours;
- alterations in immunological competence. There is evidence that damage to the immune system may lead to an increase in the incidence of tumours during an animal study.

d) The substance acted only as a promoter, ie. caused an increase in the incidence of tumours only when administered after treatment with a known carcinogen.

e) The occurrence of tumours can be reliably attributed to changes in the metabolism of the substance due to metabolic over-loading at high doses - for example, the saturation of de-toxification processes which are normally present in man and the test species.

f) The existence of evidence from an epidemiological study or studies which is adequate to contradict that derived from limited data in animal studies.

g) The demonstration of carcinogenic activity in animal studies at exposure levels which greatly exceed the likely human exposure, in the absence of evidence that the substance is genotoxic.

D. CLASSIFICATION OF MUTAGENS

In the 5th Adaptation mutagenic substances are classified into three Categories as follows :

"Category 1. Substances known to be mutagenic to man. There is sufficient evidence to establish a causal association between human exposure to a substance and heritable genetic damage.

Category 2. Substances which should be regarded as if they are mutagenic to man. There is sufficient evidence to provide a strong presumption that human exposure to the substance may result in the development of heritable genetic damage, generally on the basis of appropriate animal studies.

Category 3. Substances which cause concern for man owing to possible mutagenic effects but in respect of which the available evidence does not satisfactorily demonstrate heritable genetic damage. There is evidence from appropriate mutagenicity studies, but this is insufficient to place a substance in Category 2."

Although mutagenicity tests are often conducted for carcinogenicity screening, mutagenicity is a toxicological property in its own right. In the following sections the rationale for placing mutagenic substances in one or other of the above Categories is discussed.

1. Category 1 : Substances Known to be Mutagenic to Man

   Evidence for inclusion in this Category can come only from detailed observation of exposed human populations. Clinical and epidemiological studies would be required to demonstrate a causal relationship between exposure to a substance and the occurrence of heritable effects (germ-cell mutations). However, no human mutagens are known and it is extremely difficult to gain reliable information from studies on humans for the following reasons :

   a) there is a large variation within the human gene pool;
   b) the number of progeny is small;
   c) the generation time is long;
   d) not all mutations produce detectable phenotypic variations;
   e) only dominant, and some sex-linked, recessive mutations would be observed in the first generation;

f) many mutations are likely to be lethal at an early stage of development of the embryo.

2. Category 2 : Substances Which Should be Regarded as if They Are Mutagenic to Man

Evidence for the inclusion of a substance in this Category would come from appropriate animal studies and other relevant information. The studies should provide sufficient evidence that exposure to a substance may result in heritable genetic damage.

2.1 Animal studies

Mutagenic activity may be detected as gene mutations, chromosome or genome mutations (aneuploidy) or through the detection of DNA damage in the form of sister chromatid exchanges, DNA-breaks or DNA repair. Animal studies to detect these end-points are available. Genetic damage may be induced in either somatic or germ cells. The latter are most important for assessing the risk of heritable genetic damage to humans.

2.1.1 Mammalian tests which directly demonstrate heritable genetic damage.

Some mammalian tests which are capable of directly demonstrating the induction of heritable genetic damage in mammals have been developed, e.g. the specific locus test or the heritable translocation test. These tests were originally used to assess the effects of direct radiation on the gonads in animals. When used to assess the effect of chemicals they may require extremely large numbers of animals (more than 10,000 progeny) and are nevertheless inherently insensitive because of the small number of loci examined. Consequently, they are not used for routine testing.

2.1.2 Mammalian tests for mutagenic activity not having heritable effects as an end-point. Several mammalian test systems have been developed to detect different mutagenic end-points. Some detect mutations in the germ cells (e.g. germ-cell metaphase analysis or the dominant lethal test), whilst others assess the effects in somatic cells (for example, in the bone marrow, circulating lymphocytes or the liver). Indirect tests for detecting DNA-damage in germ cells are also available, e.g. the induction of unscheduled DNA-synthesis, sister chromatid exchange, or strand-breaks/fragmentation; and covalent binding to DNA.

2.2 Other relevant information

For the classification of mutagens, other relevant information may be obtained from *in vitro* tests with non-mammalian or mammalian cells. Many such tests have been developed to screen for mutagenic activity as an indication of carcinogenicity. Still other test systems enable mutagenic activity to be detected in sub-mammalian species *in vivo*.

2.2.1 Bacterial mutagenicity tests. In these tests mutagenicity is generally detected by the induction of either forward or reverse gene mutations. Bacterial tests are rapid, and because large numbers of bacteria are exposed to the test compound they do not suffer from the statistical problems of sensitivity associated with *in vivo* studies. However, bacteria may be more susceptible to the effects of mutagenic chemicals than are mammalian cells because of their low capacity for DNA repair (repair-deficient strains are often used), the absence of a nuclear membrane, and the selection of bacteria with a cell membrane easily permeable to chemicals. Bacteria are often unable to metabolise chemicals to mutagens in the way in which mammals do, and their use can lead only to the detection of the mutagenic activity of chemicals which interact directly with DNA (with a few exceptions). This problem is overcome to a certain extent by using subcellular fractions from mammalian tissues in metabolic activating systems. However, metabolism in these systems does not always reflect the *in vivo* situation since essential co-factors may be lacking and the rate and pattern of metabolism are not considered.

Host-mediated assays have been developed in which both the test bacteria and test substance are injected into mammals. The bacteria are exposed to the metabolites of the test substance in the animal, so that the metabolism and toxicokinetics of the substance are those of the intact animal. After a time the bacteria are recovered from the animal and mutants are detected and counted. The value of such tests may be limited by the fact that reactive metabolites have to diffuse to the bacteria, and the tests cannot be regarded as equivalent to mammalian *in vivo* tests. Thus, bacterial tests provide a valuable indication of mutagenic potential, but cannot be used on their own for classification purposes.

2.2.2 Mutation assays with mammalian cells in culture. In vitro mammalian test systems have been developed to enable all three mutagenic end-points (eg. gene, chromosome, and genome mutations) to be detected. Other systems, for example the induction of DNA-repair or sister-chromatid exchanges, detect genetic damage indirectly. However, cultured mammalian cells are often unable to metabolise chemicals into their mutagenic metabolites and so require externally-added activating systems. Such test systems provide a valuable indication of mutagenic potential, but results from them cannot be used on their own for classification purposes.

2.2.3 Studies on exposed human populations. The induction of genetic damage in exposed human populations may be indicated by an increased incidence of chromosome aberrations or sister-chromatid exchanges in circulating blood cells (eg. lymphocytes). In general, the incidence found varies markedly both from individual to individual and in samples taken from a single individual. Thus valid information is difficult to obtain. Reliable evidence for a causal relationship between exposure to a substance and an increased incidence of genetic damage in circulating cells must be taken as a demonstration of somatic mutation in humans. However, in the absence of evidence of an effect on germ cells it cannot be assumed that the substance will produce heritable genetic damage.

2.2.4 Metabolic studies. Most chemicals require metabolism to reactive intermediates before interaction with DNA occurs, thereby inducing the mutations. Metabolic studies may provide evidence for the formation of reactive metabolites but this does not prove that the substance is mutagenic. Consequently, metabolic studies cannot be used as a basis for classification.

2.3 Criteria for classification in Category 2

For the classification of a substance as a Category 2 mutagen either of the following requirements must be met :

a) there should be positive findings in at least one in vivo germ-cell mutagenicity test in mammals, i.e. from the specific locus, heritable translocation or dominant lethal tests, or from germ-cell metaphase analysis :

or

b) there should be positive findings in at least one in vivo somatic cell mutagenicity test in mammals or evidence of somatic cell mutagenesis in humans, plus sufficient evidence that the substance in its active form interacts with germ-cell DNA following systemic administration by an appropriate route.

The somatic cell mutagenicity tests may comprise i) in vivo cytogenetic studies, i.e. bone-marrow metaphase analysis, micronucleus tests or lymphocyte chromosomal analysis, or ii) the mouse coat-colour spot test.

Whereas the principle of the approach in b), above, is acceptable, its implementation in practice may pose problems with regard to the definition of what is sufficient evidence that the substance interacts with germ-cell DNA. However, possible criteria for such an interaction would be the detection of any one of the effects listed below after systemic administration of the substance by an appropriate route :

i) the induction of unscheduled DNA-synthesis (DNA-repair) in germ-cells;

ii) covalent binding of the substance to germ-cell DNA resulting in extensive adduct formation ;

iii) the induction of sister chromatid exchanges in germ-cells;

iv) the induction of DNA-damage (strand breaks, fragmentation) in germ-cells.

Although there are tests for determining the above effects, they are at present poorly validated. Results are available only with a small number of chemicals from a few laboratories. Great care must be taken in interpreting the results of such tests, and the interpretation should be on a case by case basis.

3. Category 3 : Substances Which Cause Concern Because of Possible Mutagenic Effects

Substances would be included in this Category when there is sufficient evidence of mutagenic activity in somatic cells but insufficient evidence that the substance interacts with mammalian germ-cell DNA.

3.1 Criteria for a Category 3 classification

For the classification of a substance as a Category 3 mutagen the following are required :

a) positive findings in at least one in vivo somatic cell mutagenicity assay in mammals without sufficient evidence that the substance in its active form interacts with mammalian germ-cell DNA ;

or

b) other evidence of genotoxic activity from in vivo studies in mammals (e.g. the induction of sister chromatid exchanges, unscheduled DNA synthesis, or DNA breakage) accompanied by evidence from tests on cultured mammalian cells or Drosophila which indicates that the substance induces point-mutations or chromosomal aberrations.

If positive findings are available only from in vitro tests, an appropriate in vivo study should performed before any proposal for classification or non-classification is made.

4. Circumstances Leading to Non-classification

In the absence of evidence for mutagenicity a substance is clearly not classified. A substance may not be placed into any of the above Categories if any of the following apply :

a) evidence of mutagenicity is available only from in vitro tests. Such findings should lead to in vivo testing;

or b) direct evidence indicates that the substance or any of its metabolites do not interact with mammalian germ-cell DNA.

E. CLASSIFICATION OF TERATOGENS

In the 5th Adaptation, teratogenic substances are required to be classified in two Categories:

"Category 1. Substances known to be teratogenic to man. There is sufficient evidence to establish a causal association between human exposure to a substance and subsequent non-heritable birth defects in offspring.

Category 2. Substances which should be regarded as if they are teratogenic to man. There is sufficient evidence to provide a strong presumption that human exposure to the substance may result in non-heritable birth defects in offspring, generally on the basis of :

- appropriate animal studies,
- other relevant information."

1. Introductory Remarks

Teratogenicity is only one manifestation of the broad spectrum of adverse effects which substances may produce on the developing embryo and foetus. Other manifestations, eg. embryolethality, are not considered in the above classification. In the past decade the scope of the term "teratogenicity" has broadened. Whereas it was previously understood only as the induction of morphological anomalies, there is nowadays an increasing tendency to include also functional anomalies. This is reflected in the OECD guidelines for teratogenicity studies where "teratogenicity" is defined as "the property of a chemical which causes permament structural or functional abnormalities during the period of embryonic development". However, none of the national or international guidelines for such studies deals with the recording of functional anomalies. Furthermore, no sufficiently-validated test procedures for the post-natal detection of such anomalies in single- or multi-generation studies are available. In view of this, classification as a teratogen from the results of animal experiments can be based only on the induction of morphological abnormalities.

No effect-levels exist for teratogens and thus the likely level of human exposure must be taken into account when considering the classification of a substance as a teratogen.

For more detailed information on the various aspects of teratogenicity see ECETOC (1983), Wilson and Frazer (1977-8) and Kimmel and Buelke-Sam (1981).

In the following sections the rationale for placing substances in one or the other of the above Categories is discussed.

2. Category 1 : Substances Known to be Teratogenic to Man

Substances to be put in this Category are those for which there is sufficient evidence that they cause structural anomalies in the progeny of women exposed to them during the critical period of pregnancy. Such

evidence can result only from well-conducted epidemiological studies or an accumulation of well-documented case studies. A causal relationship between the exposure of women to a substance during the critical period of gestation and the occurrence of birth defects in their offspring is indicated by one or more of the following observations :

i) a marked increase in the frequency of abnormal infants in women exposed to a substance. The significance of this would be emphasised if the frequency were higher at higher exposures;

ii) a sudden increase in the frequency of a specific anomaly in the offspring;

iii) the occurrence of a rare type of malformation in the offspring.

In each case the possibility that the effect could have been caused by confounding factors such as nutritional state, alcohol abuse and exposure to other chemicals, etc. must be excluded. A specific problem is that in many countries reliable background data on congenital malformations in the general population do not exist.

3. Category 2 : Substances Which Should be Regarded as if They Are Teratogenic to Man.

Substances to be put in this Category are those for which the results of appropriate animal studies and other relevant information provide evidence that there is a strong presumption of a teratogenic risk.

3.1 Data from animal studies

The design, conduct and results of the animal studies leading to a Category 2 classification must be appropriate.

3.1.1 Appropriateness of design and conduct of a study. Past and current studies differ little in their design, those carried out nowadays normally being performed in accordance with recognised guidelines and GLP. Studies carried out in accordance with such guidelines are generally regarded as appropriate, as are the data so produced. The acceptability of teratogenicity studies carried out before such guidelines were published has to be judged on a case by case basis, but generally each of the following minimum requirements should be fulfilled.

a) The test substance should have been clearly specified, and administered at defined dose levels.
b) The studies must have been conducted on whole mammals since embryonic development within the maternal body is necessary.
c) The route of administration of the test substance to the mothers should have been relevant to the route of human exposure (normally oral, dermal or by inhalation).
d) The number of animals used per dose level should have been high enough to establish whether the response was treatment-related. The OECD recommend 20 pregnant females per dose level for rodents and 12 for lagomorphs.
e) The test substance should have been administered at least during organogenesis, the most susceptible period of gestation.
f) Foetal examination should have been competently carried out by generally-accepted methods.

3.1.2 Appropriateness of the results of a study. In order to conclude that a substance is teratogenic in an animal model, the results must meet the following essential criteria :
a) the frequency of the anomalies recorded must be statistically and biologically significant in the species and strain used ;
b) a clear dose-response relationship for the teratogenic effects produced should normally have been observed;
c) the teratogenic effects must have occurred at at least one dose level which did not cause overt maternal toxicity.

3.1.3 Criteria for Classification in Category 2. The occurrence of teratogenic effects meeting the above criteria in one mammalian species is regarded as sufficient for a Category 2 classification of the substance concerned, unless the relevance to man is questionable - see section 4, below.

3.1.4 Other relevant information. There is often little further relevant information available other than that from in vivo animal experiments. In most cases, classification of a substance in Category 2 has to be made on the basis of information derived from a classical teratology study or studies. Although biochemical and metabolic data may assist in interpreting the results of certain mammalian studies, it would only rarely be directly useful for classification. The results from in vitro

and non-mammalian tests may be useful for screening purposes but not for classification, since their effectiveness has not yet been sufficiently validated. Structure-activity relationships are not at present sufficiently well understood to provide a basis for classifying a substance as a teratogen, as it is known that small structural changes in a molecule can be decisive in determining whether it is a teratogen.

4. Circumstances Leading to Non-classification

Substances which do not meet the above-described criteria for inclusion in either Categories 1 or 2 should not be classified. In addition, the following circumstances may cast adequate doubt on the relevance of the study results to man, and so lead to non-classification:

a) the occurrence of teratogenic abnormalities only in the presence of overt maternal toxicity;

b) the occurrence of teratogenic effects in animal experiments only at dose levels grossly in excess of the likely human exposure ;

c) the induction of teratogenic effects in animal experiments with exposure routes which are not relevant to those of man;

d) an increase, in the offspring, of the frequency of structural changes which are simply variations within the biological norm of the species;

e) the existence of adequate evidence that the metabolite responsible for the teratogenic effect is not produced in man.

F. BIBLIOGRAPHY

- ECETOC (1982). Risk Assessment of Occupational Chemical Carcinogens, Monograph no.3, Jan. 1982.
- ECETOC (1984). The EEC Sixth Amendment : A Guide to Risk Evaluation for Effects on Human Health. Technical Report no.14, 20.3.84.
- ECETOC (1983). Identification and Assessment of the Effects of Chemicals on Reproduction and Development (Reproductive Toxicology), Monograph no.5, Dec. 1983.
- IARC (1984). IARC Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Humans, Vol.33, p.15.
- Kimmel, C.A. and Buelke-Sam, J. (1981), Eds., Developmental Toxicology, Raven Press, New York.
- Wilson, J.G. and Frazer, F.C. (1977-8), Eds., Handbook of Teratogenicity, Plenum Press, New York and London.

G. MEMBERS OF TASK FORCE

Dr. J.M. Faccini (chairman)	Consultant to Rhône-Poulenc
Dr. P. Bentley	Ciba Geigy, Basel
Dr. H. Fleig	BASF, Ludwigshafen
Dr. A.J. Ingram	BP, Sunbury
Dr. R. Millischer	ATOCHEM, Paris
Dr. H. Schön	Sandoz, Basel
Dr. R.W. Stephens	Shell, London

H. MEMBERS OF ECETOC SCIENTIFIC COMMITTEE

I. PURCHASE (Chairman), Director, Central Toxicology Laboratory	ICI (Alderley Park)
M. SHARRATT, (Vice-chairman), Senior Toxicologist	BP (Sunbury)
J. BÄCKSTROM, Consultant Toxicologist	ASSOCIATION OF SWEDISH CHEMICAL INDUSTRIES (Stockholm)
B. BROECKER, Coordinator, Product-related Environmental Problems	HOECHST (Frankfurt)
L. CAILLARD, Industrial Toxicological Service	RHONE-POULENC (Paris)
H.O. ESSER, Head of Biochemistry, Agricultural Division	CIBA-GEIGY (Basel)
P. GELBKE, Head, Department of Toxicology	BASF (Ludwigshafen)
U. KORALLUS, Medical Director	BAYER (Leverkusen)
H.G. NOSLER, Head, Coord. Centre for Environmental Protection and Consumer Safety	HENKEL (Düsseldorf)
J.F. NEWMAN, Consultant to ICI	(Reading)
S. PAGLIALUNGA, Head of Industrial Toxicology	MONTEDISON (Milan)
C.L.M. POELS, Envir. Affairs Division	SHELL (den Haag)
C. DE SLOOVER, Head, Medical Department	SOLVAY (Brussels)
W.F. TORDOIR, Head of Occupational Health and Toxicology Division	SHELL (den Haag)
H. VERSCHUUREN, Head, Department of Toxicology	DOW CHEMICAL (Horgen)