

Monograph No. 24

Risk Assessment for Carcinogens

July 1996

ISSN-0773-6347-24

Brussels, July 1996
© ECETOC copyright 1996

ECETOC Monograph No. 24

© Copyright - ECETOC (European Centre for Ecotoxicology and Toxicology of Chemicals),
4 Avenue E. Van Nieuwenhuysse (Bte 6), 1160 - Brussels, Belgium.

All rights reserved. No part of this publication may be reproduced, copied, stored in a retrieval system or transmitted in any form of by any means, electronic, mechanical, photocopying, recording or otherwise without the prior written permission of the copyright holder. Applications to reproduce, store, copy or translate should be made to the Secretary General. ECETOC welcomes such applications. References to the document, its title and summary contained in data retrieval systems may be copied or abstracted.

The content of this document has been prepared and reviewed by experts on behalf of ECETOC with all possible care and from the available scientific information. It is provided for information only. ECETOC cannot accept any responsibility or liability and does not provide a warranty for any use or interpretation of the material contained in the publication.

RISK ASSESSMENT FOR CARCINOGENS

CONTENTS

LIST OF ABBREVIATIONS USED	1
GLOSSARY	2
SUMMARY	3
1. INTRODUCTION	5
2. BASIC ELEMENTS OF RISK ASSESSMENT FOR CARCINOGENS	7
2.1 INTRODUCTION	7
2.2 HAZARD IDENTIFICATION	7
2.3 DOSE- RESPONSE (EFFECTS) ASSESSMENT	9
2.4 RISK CHARACTERISATION	11
3. MATHEMATICAL MODELS	13
3.1 INTRODUCTION	13
3.2 MATHEMATICAL MODELS	14
4. PROBLEMS AND LIMITATIONS	19
4.1 ESTIMATION OF RISK FROM SEVERAL MODELS	19
4.2 THE APPLICATION OF THE MAXIMUM LIKELIHOOD ESTIMATE, MULTI- STAGE AND THE WEIBULL MODELS	20
4.3 VARIATION OF DOSE RANGE AND RESPONSE CURVE	22
4.4 EFFECT OF TOP DOSE	22
4.5 RELATIONSHIP OF MAXIMUM TOLERATED DOSE AND CARCINOGENIC POTENCY	23
4.6 LIMITATIONS OF THE DATA SETS USED	24
5. DISCUSSION	25
5.1 DESIGN OF CONVENTIONAL BIOASSAYS	25
5.2 USE OF MATHEMATICAL MODELS	26
5.3 CURRENT APPROACHES TO RISK ASSESSMENT	27
5.4 IMPROVEMENTS IN RISK ASSESSMENT	28

5.5 ALTERNATIVE APPROACHES TO BASIC MATHEMATICAL MODELLING IN RISK ASSESSMENT FOR GENOTOXIC CARCINOGENS.....	30
6. RECOMMENDATIONS.....	32
APPENDIX I: CLASSIFICATION SCHEMES FOR CARCINOGENIC SUBSTANCES.....	40
BIBLIOGRAPHY.....	44
MEMBERS OF TASK FORCE.....	49
MEMBERS OF THE SCIENTIFIC COMMITTEE.....	50

LIST OF ABBREVIATIONS USED

(Terms marked with * are explained in the glossary)

ADI	Acceptable Daily Intake
LCL	Lower Confidence Limit
LMS	Linearised Multi-Stage model
LOAEL	Lowest Observed Adverse Effect Level
MLE *	Maximum Likelihood Estimate*
MTD	Maximum Tolerated Dose
MVK	the Moolgavkar-Venzon-Knudsen model
NCI	National Cancer Institute (USA)
NOAEL	No Observed Adverse Effect Level
NTP	National Toxicology Programme (USA)
PB-PK	Physiologically-Based Pharmacokinetic
RSD *	Risk Specific Dose*
TDI	Tolerable Daily Intake
UCL *	Upper Confidence Limit*
US-EPA	Environmental Protection Agency (USA)
VSD *	Virtually Safe Dose*

GLOSSARY

q_i

Slope of the dose-response curve of the MLE for a specified risk.

q_i^*

The 95% upper bound estimate of the linearised slope of the dose-response curve in the low dose region determined by the Linearised Multi-Stage model (LMS).

RSD (Risk Specific Dose)

The dose corresponding to a specified level of risk (usually refers to a risk of 1 in 10^6).

MLE (Maximum Likelihood Estimate)

A statistical best estimate of the value of a parameter from a given data set. In this report the central or point estimate of a RSD.

UCL (Upper Confidence Limit)

This is a calculated value which attempts to reflect the variability inherent in the original data. The value expresses a risk with a specified confidence prediction that the true risk will not exceed this value.

95% UCL

The value for 95% confidence that the true value does not exceed that stated. The true value may be much lower.

SUMMARY

Assessment of the risk of exposure to chemical carcinogens involves: 1) hazard identification (intrinsic toxicological properties), 2) assessment of the dose response relationship, 3) determination of the (potential for) exposure, and 4) characterisation of the risk. Although there is considerable uniformity between regulatory authorities in the methods for hazard identification, the remainder of the process differs substantially, largely due to the divergent concepts applied to assessment of dose response (and therefore risk characterisation). In the USA carcinogenic risk assessment may involve mathematical modelling of experimental data, with extrapolation below the range of observations, leading to an upper bound probability estimate of human cancer risk at a given exposure level. Invariably this approach assumes a lack of a threshold dose for cancer induction and a linear extrapolation to low doses. Carcinogenic risk assessment in Europe is usually more qualitative including a thorough investigation of data taking into account the weight of all available evidence.

With the introduction of risk assessment (OECD, EU, USA) in the regulatory process for individual substances a closer examination of existing methodology, in particular the use of mathematical models for dose-response extrapolation for carcinogenic risk assessment, is timely. This report summarises the basic elements of risk assessment, describes the mathematical models available for evaluation of carcinogenic risk, their use and interpretation.

The uncertainties surrounding the use of mathematical models are explored. A number of equations are available which fit empirically the data set derived from the bioassay. For a given set of data these can give risk estimates differing by several orders of magnitude. It is currently not possible to determine which is the most appropriate. The linearised multi-stage model (LMS) is most commonly used as its estimation of the risk is likely to be more conservative than the actual risk. To determine an upper bound for the probable risk, some authorities, notably the US Environmental Protection Agency (US-EPA) have used the 95% upper confidence limit (95% UCL¹) derived from this model rather than the maximum likelihood estimate (MLE).

Illustrative examples show that:

- The UCL is relatively insensitive to changes in the experimental tumour incidence and therefore lacks discrimination; furthermore, these carcinogenic risk assessments do not discriminate between positive and negative data;
- The MLE, while distinguishing between negative and positive data, can be oversensitive to small changes in tumour number.

¹ For definitions and abbreviations used see beginning of the report.

- Both values show a greater dependence on the value of the highest dose tested rather than on the dose response within a data set. In part these findings reflect the limitations of the data set derived from the carcinogen bioassay. For modelling, a satisfactory empirical fitting of equations to at least three bioassay data points is needed. The constraints on the current bioassay design mean that, at most, two positive dose levels are included. Frequently only one of these shows a significantly-increased tumour incidence. This is compounded by the lack of any theoretical support for the basis of extrapolation to doses several orders of magnitude below the experimentally tested range.

The limitations in the use of pre-designated equations for carcinogenic risk assessment have been identified. They indicate that the approach is generally unsuitable and may not be used in a mechanical and uncritical manner.

Mathematical models taking into account the known biology of cancer induction would be more useful and could be applied on a case-by-case basis. Approaches currently under development incorporate more-sophisticated parameters than numerical tumour incidence and include physiologically-based pharmacokinetic (PB-PK) and biologically-based models. The latter require further development.

When conducting risk assessment for carcinogens it is recommended that a thorough investigation of all available evidence is made on a case-by-case basis and that suitable assessment factors are used.

1. INTRODUCTION

Risk assessment for carcinogens involves four stages. The first of these is the identification of the hazard, followed by the assessment of a dose-response relationship, the exposure assessment and the risk characterisation. The current terminology in use for hazard and risk assessment varies according to the regulatory authority concerned, which is responsible for some of the confusion which surrounds the process. The nomenclature of the processes used by various regulatory authorities can be divided into three basic steps and is summarised as follows:

Risk assessment stage	I	II + III	IV
EU (EEC, 1993)	Hazard identification	Dose-response (effect) assessment and exposure assessment	Risk characterisation
USA (US-EPA, 1986)	Hazard identification	Dose-response assessment and exposure assessment	Risk characterisation
UK (COSHH, 1988)	Hazard identification	Exposure assessment and risk assessment	Risk assessment
UK (CIA, 1991)	Hazard identification	Qualitative risk assessment	Quantitative risk assessment
GERMANY (MAK, 1991)	Hazard identification		Risk assessment

In this document the terminology used by the EU (EEC, 1993), namely hazard identification, dose-response (effect) assessment, exposure assessment and risk characterisation will be used unless referring specifically to processes used elsewhere.

The process for estimating the risk to man of exposure to chemicals identified as carcinogens varies between different regulatory authorities in the way that the dose-response assessment and the risk characterisation is carried out. This in part reflects the different objectives of the processes in different countries. In the EU, currently most clearly exemplified by the process used in the UK, the overall objective can be described as a best effort to provide an estimation of the actual risk to man given all the uncertainties concerned. By contrast, the regulatory agencies in the United States seek to establish a conservative upper bound to the risk so that the true risk is certainly lower than the calculated value and may be as low as zero. Countries also differ in the extent to which they apply mathematical equations as surrogate models for dose-response relationships in the risk assessment of chemicals identified as carcinogens. In the USA, mathematical concepts are being applied to both genotoxic and non-genotoxic carcinogens, which are in any case rarely considered separately. The US-EPA has embarked on a

process to update guidelines taking full account of current scientific knowledge. This process has proceeded in parallel with the deliberations of this Task Force. In Europe, risk assessment has not generally involved mathematical modelling, but rather considers the data in a qualitative manner (UK DOH, 1991).

With this in mind ECETOC formed a Task Force with the following terms of reference:

- to describe the methods available together with their applicability for the mathematical modelling of carcinogenic risk from toxicological data;
- to evaluate the scientific and technical merits, using examples where available, of these methods;
- to recommend how risk assessment may be improved, e.g. by the incorporation of additional data and the use of improved models.

This report summarises the basic elements of risk assessment of carcinogenic substances (Section 2), defines the mathematical models in current use and those under development (Section 3) and illustrates the interpretative difficulties which can arise from application of the equations commonly used (Section 4). Discussion and conclusions on the use of quantitative risk analysis (Section 5) and the Task Force recommendations for further investigation (Section 6) complete the report.

2. BASIC ELEMENTS OF RISK ASSESSMENT FOR CARCINOGENS

2.1 INTRODUCTION

In general, the risk assessment process, as defined by the EU (EEC, 1993), can be divided into four identifiable steps. This may involve the application of qualitative and quantitative procedures and sometimes, but not always, requires interpretation of complex biological procedures (such as pharmacokinetic modelling) or rigorous mathematical analysis. The four steps and their definitions are as follows (EU Regulation (EEC) No 1488/94):

Hazard Identification

Identification of the adverse effects which a substance has an inherent capacity to cause.

Dose (Concentration)- Response (Effects) Assessment

Estimation of the relationship between dose of, or level of exposure to a substance, and the incidence and severity of an effect.

Exposure Assessment

Determination of the emissions, pathways and rates of movement of a substance and its transformation or degradation, in order to estimate the concentrations/doses to which human populations or environmental spheres (water, soil and air) are or may be exposed.

Risk Characterisation

Estimation of the incidence and severity of the adverse effects likely to occur in a human population or environmental sphere (water, soil and air) due to actual or predicted exposure to a substance, and may include 'risk estimation', i.e. the quantification of that likelihood.

2.2 HAZARD IDENTIFICATION

When identifying the intrinsic carcinogenicity of a chemical there are several key elements in a data set:

- epidemiology
- animal carcinogenicity, (from life-time carcinogenicity bioassay)
- mechanism of carcinogenic action
- genotoxic activity and DNA adduct formation, (from short-term *in vivo* and *in vitro* tests)
- pharmacokinetic activity,
- structure activity relationships.

Analysis of structure activity relationships (SAR) relies on previous knowledge of chemicals of similar structure known to be carcinogenic to make a prediction of the carcinogenicity of a chemical of known structure but whose carcinogenicity is unknown. Clearly this application can be of maximum benefit when establishing priorities for testing within a series of closely aligned chemical components or structural analogues. SAR alone cannot be used as a surrogate for toxicity testing at the current state of development (ECETOC, 1986).

Hazard identification applies a weight of evidence approach, using the data listed above, to classify chemicals into groups. Various classification schemes are used, all are based on the sufficiency of the evidence of carcinogenicity in animals or in man. The best known schemes for categorising carcinogenic chemicals are those of IARC (1987), EEC (1991) and US-EPA (1986). These classification schemes are summarised in tabular form in Appendix 1. The attainment of a specified classification level is generally the trigger to proceed to the next step in the risk assessment.

The IARC scheme was not produced for regulatory purposes but has been extensively used in developing regulatory guidelines. The EPA scheme has been used as a decision point for entering quantitative risk assessment. Some chemicals classified as category C and all category A and B carcinogens enter quantitative risk assessment. The EU scheme is currently used to assign a hazard (classification) symbol and risk phrases to the chemical label.

A common criticism of classification schemes is that they generally give precedence to positive results over negative results even if the former are of poor quality and the latter relatively extensive. Current schemes may not fully accommodate evidence that some mechanisms of carcinogenicity are species specific and therefore the carcinogenicity data are not relevant for human hazard assessment. Recently, recognition of these difficulties has led to proposals for alternative classification schemes such as that proposed by Ashby *et al* (1991); Scheme 4 in Appendix I. Also the US-EPA has embarked on a process to update guidelines to take full account of the current state of scientific knowledge. This process has proceeded in parallel with the deliberations of this Task Force.

Examples of mechanistic data which may be expected to alter the classification of some chemicals, include tumours produced specifically in the male rat kidney following α_2 -microglobulin accumulation in tubular cells (Swenberg *et al*, 1989; US EPA, 1991 a,b) and thyroid follicular cell carcinomas produced in rodents after exposure to substances capable of inducing thyroid gland enlargement (goitrogens) (Paynter *et al*, 1988 ; Atterwill *et al*, 1992).

It is generally recognised that carcinogens may be divided into two classes, genotoxic and non-genotoxic. Genotoxic carcinogens act either directly or indirectly on DNA to modify its base structure, which is assumed to be the basis of their carcinogenicity. Non-genotoxic chemical carcinogens are believed to

exert their carcinogenic effects through processes which do not involve direct binding of the chemical or its metabolites to DNA but rather mimic the effect of some natural signal transduction factors or cause cell proliferation and therefore perturb normal cell growth and differentiation (Schuller, 1991; Green, 1991). Currently, genotoxic carcinogens, for regulatory purposes, are assumed to have no threshold dose for the expression of carcinogenesis, i.e. exposure to even very low doses of a genotoxic carcinogen is assumed to present a small but finite risk of tumour formation. As non-genotoxic carcinogens are generally considered to interfere with processes under homeostatic control, they are expected to show a threshold dose below which there will be no observed tumour incidence. Thus identification of genotoxic potential is of particular importance in the definition of a chemical carcinogen as this characterisation may determine the way in which the risk assessment is conducted.

2.3 DOSE- RESPONSE (EFFECTS) ASSESSMENT

Epidemiological studies rarely discriminate between effects at different exposure levels with much precision. Frequently the available information on exposures is rudimentary and even when more-detailed exposure assessments have been made their relevance to an individual is subject to a high degree of uncertainty. Thus, dose-response assessments generally rely heavily on surrogate data obtained from animal experiments.

The dose or level of exposure that causes an effect in an animal study and the likely dose causing a similar effect in humans is a key consideration in risk assessment. In the absence of further information, the administered dose is usually used in assessing risk. However, this introduces a considerable element of uncertainty. Four aspects of dose should be considered in risk estimation: administered dose, absorbed dose, delivered dose to the target organ and the biologically-effective dose. An accurate knowledge of the relationships between these doses leads to greater precision in risk assessment. Route of administration can have a particular influence. For assessment of risk, data from experiments using routes of possible human exposure (dietary, inhalation or dermal routes in the majority of cases) are more appropriate than non-physiological routes such as intravenous, intraperitoneal or intrapleural injections which may nevertheless have value in determining intrinsic hazard.

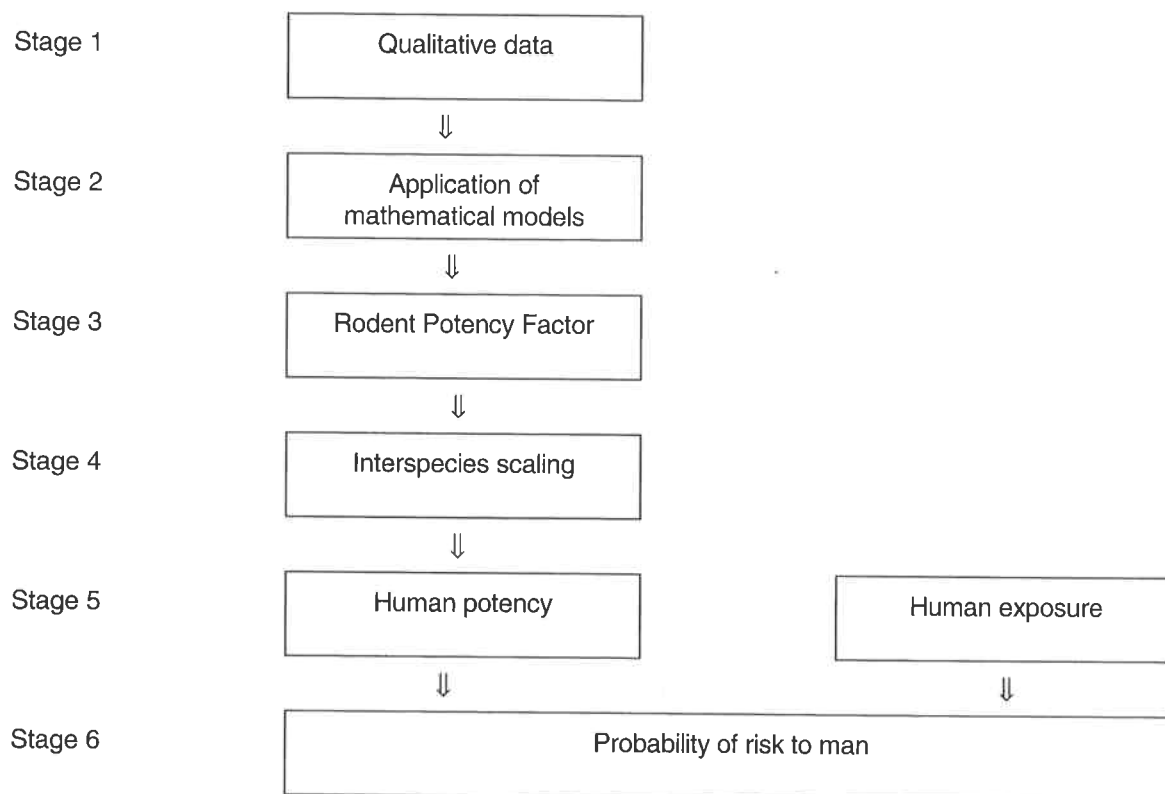
Each chemical administered will have a particular absorption rate and in most cases the chemical will undergo metabolic activation or deactivation, resulting in a delivered dose of parent compound or active metabolite to a target tissue. Metabolic and pharmacokinetic studies may help to refine estimates of the absorbed or delivered dose by defining rates of absorption, distribution, metabolism and excretion of the chemical. For precise risk assessment it is preferable to be able to use the dose of a chemical, or where appropriate its active metabolite reaching the target tissue, in the dose response assessment. This involves some knowledge of the comparative pharmacokinetics of the chemical in different species in order to estimate the delivered dose to man. Physiologically-based pharmacokinetic (PB-PK) modelling

(Connelly and Andersen, 1991), can be used to account for physiological and metabolic differences between animals and man. Such models assume that biological systems effectively consist of a number of physiological compartments and apply information on the anatomy and physiology of the test animal, the solubility of the agent in various organs (to estimate its presence) and the metabolic profile to derive a dose delivered at the target tissue. The data often come from animal studies using radio-labelled chemicals. PB-PK models have been used to refine a number of risk assessments including methylene chloride (Andersen *et al*, 1987; ECETOC, 1988, 1989), trichloroethylene (Bogen, 1988), ethylene dichloride (D'Souza *et al*, 1987), and carbon tetrachloride (Paustenbach *et al*, 1987).

When a best estimate of dose has been selected, the relationship between this dose and the established effect is assessed. For many toxicological endpoints (e.g. chronic toxicity, reproductive toxicity) an experimental No Observed Adverse Effect Level (NOAEL) or Lowest Observed Adverse Effect Level (LOAEL) is identified and used as a basis for extrapolation. In general in the EU this approach is adopted only for non-genotoxic carcinogens where a threshold effect is accepted. For genotoxic carcinogens, the method of assessment shows rather more variation. Several (European) countries rely on expert judgement, considering all available evidence (including mechanistic data), as to the likely occurrence of the carcinogenic effect in man given sufficient exposure. Other countries (USA, Denmark, Netherlands) use mathematical equations to extrapolate from the doses used in animal studies to estimate a dose that gives a predicted level of human response (Figure 1). Inevitably such extrapolations extend beyond the experimental data points.

Several different mathematical models can be fitted to the dose response data (subsequently described in Section 3). There is usually no clear evidence on which to select a specific model, and none of the equations can be given special preference. Currently the linear multi-stage model is most often used. The output of the equation is an estimate of the risk per unit dose in the test species (rodents). This is almost always the 95% upper confidence limit on risk (95% UCL or q_1^*) although in some cases a maximum likelihood estimate (MLE or q_1) may be used. US-EPA uses the linear multi-stage model as the default procedure in carcinogenic risk assessment.

Figure 1: Risk Assessment Scheme using Mathematical Modelling as Applied by the U.S. EPA



To account for differences in metabolising capacity between man and test animals the US regulatory authorities incorporate an additional safety factor to convert the rodent potency factor to a human potency factor (US-EPA, 1992). This factor is based in part on the assumption that rodents can detoxify or eliminate xenobiotics much faster than man resulting in a decrease in the effective exposure to the substance. Though this assumption may be true for ultimate carcinogens, the opposite could apply to substances which show carcinogenicity only after metabolism. The carcinogenic risk may be better assessed by determining the concentration of the ultimate carcinogen at the site of action by means of PB-PK modelling.

2.4 RISK CHARACTERISATION

Once the nature of the hazard has been established, its relevance to man determined and the dose response relationship assessed, then the human risk at probable exposure levels or conversely the level of exposure which could be considered acceptable may be determined. This process can also include an estimate of the probability of cancer under defined conditions of exposure. In arriving at these estimates

due consideration must be given to the nature of the hazard and also the frequency and duration of exposure.

For general toxicity endpoints and non-genotoxic carcinogens (where a NOAEL or LOAEL is available from an animal study) a safety factor is usually applied to determine an acceptable level of exposure where no adverse effects are expected in humans (Johannsen, 1990). The concepts of Acceptable Daily Intake (ADI) and Tolerable Daily Intake (TDI) have been used for assessing food additives and pesticides (WHO, 1987, 1990) to define a dose "without appreciable risk". This is normally derived by dividing the animal-derived NOAEL by a safety factor (alternatively called assessment factor or uncertainty factor) which has a default value of 100. More precise safety factors can be applied on a case-by-case basis (Renwick, 1991; ECETOC, 1995).

For genotoxic carcinogens, in the UK and most European countries, an expert judgement is made on a case by case basis considering all the available evidence, to predict an exposure level which is unlikely to lead to an increased cancer incidence in man. Subsequently competent authorities seek to reduce exposure as far below this level as reasonably practicable, or eliminate exposure completely in co-operation with the relevant industry.

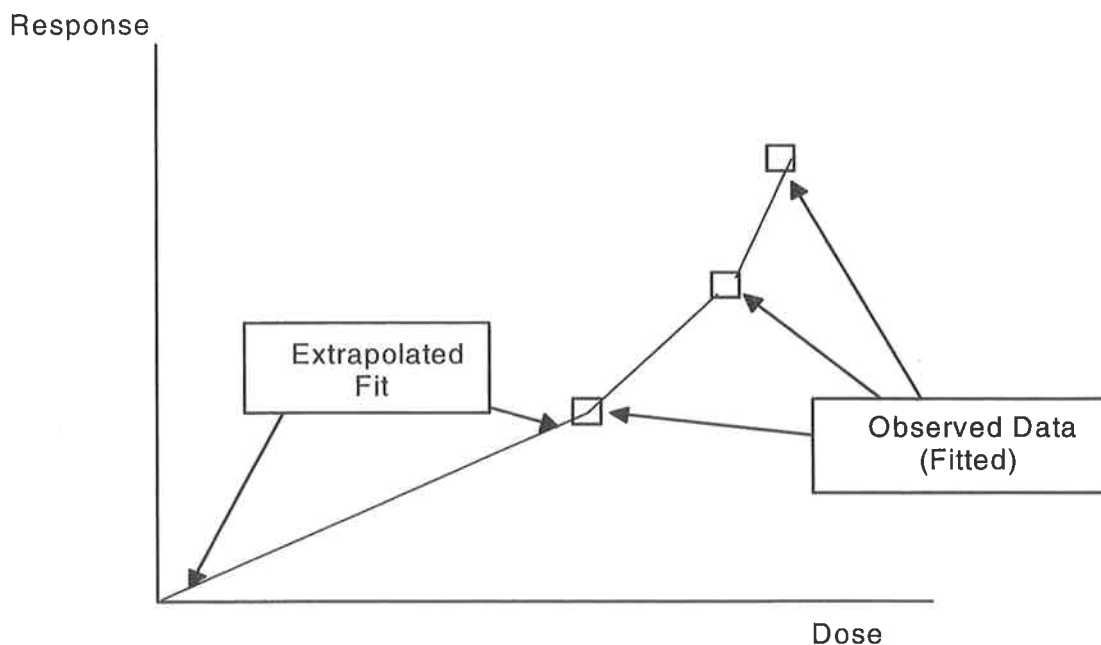
Where a mathematical model has been used to calculate a rodent or a human potency figure the probability of risk to man is calculated by comparison (multiplication) with a given human exposure level. The calculated risk is expressed as the dose associated with a predefined increase in lifetime risk of developing cancer, the Risk Specific Dose (RSD). The dose calculated to give negligible increase in risk, such as 1 in 10^6 in a lifetime, has been called the virtually safe dose (VSD) by some US regulatory agencies.

3. MATHEMATICAL MODELS

3.1 INTRODUCTION

Quantitative risk assessment involves the fitting of mathematical functions to some measure of tumour incidence, e.g. tumours observed in long-term carcinogenicity studies. These functions are based upon the mathematical models described in more detail below. The mathematical functions are then extended or extrapolated to doses far below those used in the experiment (Figure 2).

Figure 2: Diagrammatic Representation of the Use of Mathematical Models to Fit Observed Data and Extrapolation to Low Dose

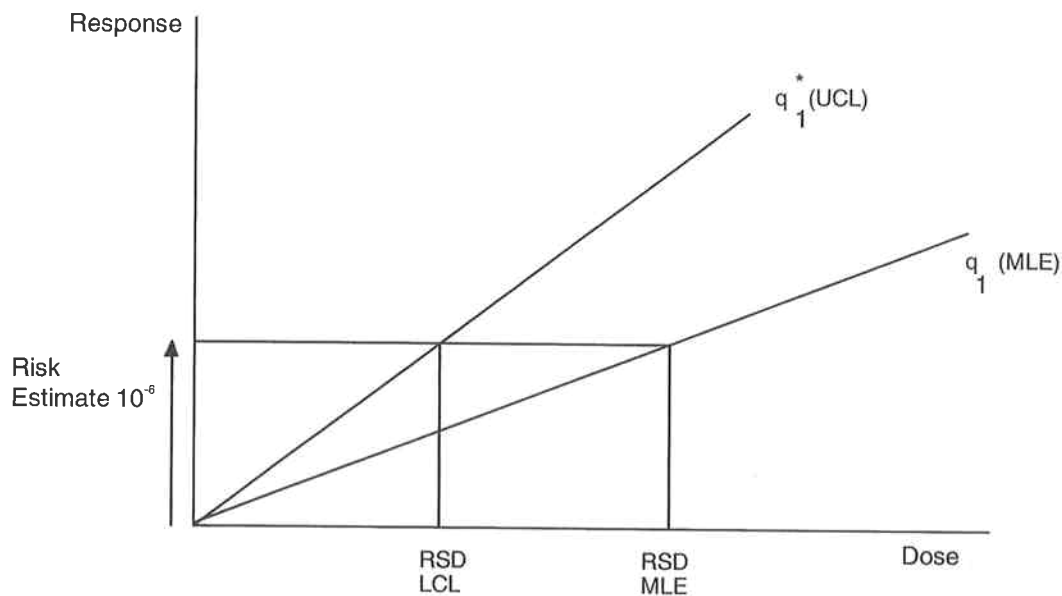


Often the values of the Risk Specific Dose (RSD) are provided as the "best" or maximum likelihood estimate (MLE) and the lower confidence limit of the dose (LCL). The difference between the two values is described below.

The maximum likelihood estimate (MLE) can be considered the central or point estimate for the parameter and the 'true' parameter value is equally likely to be greater or smaller than the MLE estimate. The precision of the MLE can be indicated statistically by the provision of a measure of the range or spread of possible estimates. This is generally expressed as a confidence interval which

gives the probability that the true value lies within the expressed range. Typically a 95% confidence limit may be quoted.

Figure 3: Diagram illustrating the use of UCL and MLE q_1 values to produce LCL and MLE of the RSD for a risk of 10^{-6}



In the case of low dose extrapolation using the linearised multi-stage model the 95% UCL of the parameter q_1 , (q_1^*) is usually quoted. One potentially confusing feature of the terminology is that when q_1^* is used to produce an estimated dose for a specified risk (say, 1×10^{-6}), the RSD obtained is referred to as the LCL RSD. The concept is illustrated in Figure 3.

3.2 MATHEMATICAL MODELS

Mathematical models are categorised loosely on a basis of their underlying statistical assumptions. These categories are termed linear, mechanistic, tolerance distribution, time-to-tumour and biologically motivated. The division between the models is somewhat arbitrary as there is considerable overlap (Table 1). Although these models claim to reflect the underlying biology in their designs, experience has shown that they represent gross over-simplification with the possible exception of the biologically motivated Moolgavkar-Venzon-Knudson (MVK) model. The mathematical models reviewed have been described in detail elsewhere (Johannsen, 1990; Paustenbach *et al*, 1990).

Table 1: List of Models Used for Carcinogen Risk Assessment

Category	Model
Stochastic or Mechanistic, and/or Linearised	One-hit Multi-hit Multi-stage Linearised Multi-stage
Tolerance distribution	Weibull Logit Probit (Mantel-Bryan)
Time-to-tumour	Log-normal distribution Weibull distribution Hartley-Sielken Armitage-Doll Multi-stage (adapted)
Biologically Motivated	Moolgavkar-Venzon-Knudson

3.2.1 Stochastic or Mechanistic Models

The one-hit model is based on the theory that a single 'hit' (i.e. DNA damage or binding to a receptor) can initiate an irreversible series of events leading to cancer and the probability of a 'hit' is directly proportional to the chemical concentration. The low-dose region of the model approximates a linear relationship. An extension of the one-hit model is the multi-hit model which assumes that multiple hits are required to initiate cancer (Hanes and Wedel, 1985).

The most widely used (by regulatory agencies) and referenced model in the literature is the multi-stage model (Armitage and Doll, 1961); of which the linearised multi-stage model is a special case (Hanes and Wedel, 1985; Crump *et al*, 1977). This model assumes that several random hits or events are required in a specific sequence for the development of cancer. It is based on the assumptions that the transition rates between successive stages are not necessarily equal, but at least one of these transitions is linearly related to dose. Linearity at low doses is based on the argument that some level of background tumours is always present and a carcinogenic agent simply

enhances or augments this process linearly (Crump, 1984). Only a small portion of the background needs to be additive for this to be true (Hoel, 1980).

The multi-stage model is sometimes considered the most plausible based on superficial similarity between it and multi-stage biological models of cancer. However, the mathematical equation derived from the multi-stage model almost certainly over-simplifies the biological process of carcinogenesis.

3.2.2 Tolerance Distribution Models

These models assume that a population contains a distribution of individuals of different susceptibilities. The models include the log-probit (Mantel and Bryan, 1961), logit and Weibull models (Weibull, 1951; Hanes and Wedel, 1985). The log-probit model utilises a linear relationship between probits and the logarithm of dose and there is a very rapid decline towards zero response. An upper bound on risk (upper confidence limit) can be calculated. The logit model differs only slightly in the logit transformation factor which shows linear dependence on the logarithm of the dose. Both models display a sigmoid curve in the experimental range and are similar at the midpoint of the curve.

More widely used is the Weibull model which has been used extensively to predict time to failure of electrical and mechanical components. It is capable of representing threshold and concave curves and is sensitive to the shape of the dose response curve. It has the advantage of being able to incorporate a time-to-tumour function.

These models are used less often than the multi-stage model.

3.2.3 Time to Tumour Models

The models described above are quantal; the only information used is the presence or absence of a tumour in an animal by a fixed time. Complications can arise in fitting equations to such binary or dichotomous data if there is differential survival between the groups.

In a long-term rodent bioassay the animal's age at death is normally recorded. Tumours can be defined as either fatal (cause of death) or incidental (death occurred from other causes). The animal's age at death can be used as an approximation of the time to the occurrence of a fatal tumour. These data can be used in the statistical analyses outlined by Peto *et al* (1980) to detect differences in time-to-tumour incidence as well as differences in overall tumour incidence between the treated and control groups.

These data can also be used in mathematical models to attempt to improve the accuracy of extrapolation to doses below those used in the experiments. The models developed to include data on the time until the tumour was observed (called time-to-tumour or time-to-observation) are generalisations of the multi-stage and Weibull equations and are based on the probability of a tumour being observed at a specified age at a given dose (Hoel, 1982, Brown and Hoel, 1983; Krewski *et al*, 1983). The most widely quoted model is the generalisation of the multi-stage model developed by Hartley and Sielken (1977). A major limitation to the use of such models has been the large number of parameters needed which require more complex experiments than the current standard two-year studies. However, Peto *et al* (1991 a,b) and Portier *et al* (1984) concluded that the Weibull model was most appropriate to describe the statistical functions of studies with time to tumour data. Armitage (1982) has suggested that the supposed advantage of time-to-tumour models over quantal models may be overstated because there may, in practice, be little extra information associated with knowledge of the time to tumour over the simple proportion of animals with tumours.

Time-to-tumour models have not been widely validated and comparisons with other models are rare. Present evidence is that they generally offer no advantage over quantal models, after the data have been adjusted to account for differences in life-span. An exception may be when the lifetimes of the animals in a bioassay were appreciably different (e.g. when the compound caused very early deaths or in a study where all of the animals died before the end of the expected dosing period).

3.2.4 Biologically Motivated Models

These models are based on knowledge of tissue growth and cell kinetics. The Moolgavkar-Venzon-Knudson (MVK) model is the only valuable example developed. This model assumes that malignant tumours arise from a single malignant cell and that malignant transformation of a stem cell is the result of two specific rate limiting irreversible events (mutations) which occur during cell division. All such models are dependent on the 'birth' and 'death' rates of cells in different stages (normal, initiated and transformed) and therefore rely heavily upon experimental data which is difficult to obtain (Moolgavkar and Leubeck, 1990). These models represent the most plausible current approach but the data necessary to validate the models are incomplete (Andersen *et al*, 1992).

3.2.5 Other Approaches

The simplest approach to extrapolation below the experimental range is a linear projection from a dose (or upper confidence limit) in the experimental range to zero. This approach, which is often termed model-free, has a number of variations. These all concern the choice of point from which to extrapolate (Krewski *et al*, 1991). This approach generally involves fitting a mathematical function to the experimental data before extrapolating linearly from a point (or upper confidence limit) on the

fitted curve where excess risk is in the observable range. Results from this approach have yielded risk estimates which are generally close to or more conservative than those obtained using the linear multi-stage model (Krewski *et al*, 1984, 1991).

Mathematical modelling of the data within the experimental range to estimate a dose corresponding to a defined level of effect which falls within or very close to the experimental range, has been applied in other circumstances, notably reprotoxicology (US-EPA, 1991c). The dose so defined has been termed a "bench mark dose" and has been used as an alternative to an NOAEL. A "safety factor" is normally applied to this dose to establish an acceptable exposure level.

4. PROBLEMS AND LIMITATIONS

The use of mathematical modelling is an attempt to express risk in quantitative terms. In doing so it generates a number of expectations concerning both the performance of the model and the accuracy of the estimated risk, including the quality of the data base used. These expectations can be summarised as:

- The estimates should differentiate between positive, equivocal and negative data;
- The fit of the equation to the data and the subsequent calculations should distinguish between different dose spacings and slope of response;
- The estimation of risk should be independent of the top dose tested;
- The use of a 95% upper confidence limit (95% UCL or q_1^*) on the maximum likelihood estimate (MLE or q_1) should reflect and be equally as accurate as the MLE itself.

In practice there are a number of limitations to mathematical modelling which affect the realisation of these expectations. These are outlined below. To illustrate the problems encountered with modelling, and the limitations of the mathematical equations currently used for risk assessment, calculations were made using a set of hypothetical data. The most widely reported and reviewed mathematical model in current use is the linearised multi-stage (LMS). This and the Weibull model have been applied to a hypothetical data set to produce the following illustrations. The data were fitted to the models using a commercially available computer programme and the q_1 (MLE), q_1^* (95% UCL) and the related RSD's for a 1×10^{-6} in a lifetime risk were determined (Tables 2-4, see pages 34-39).²

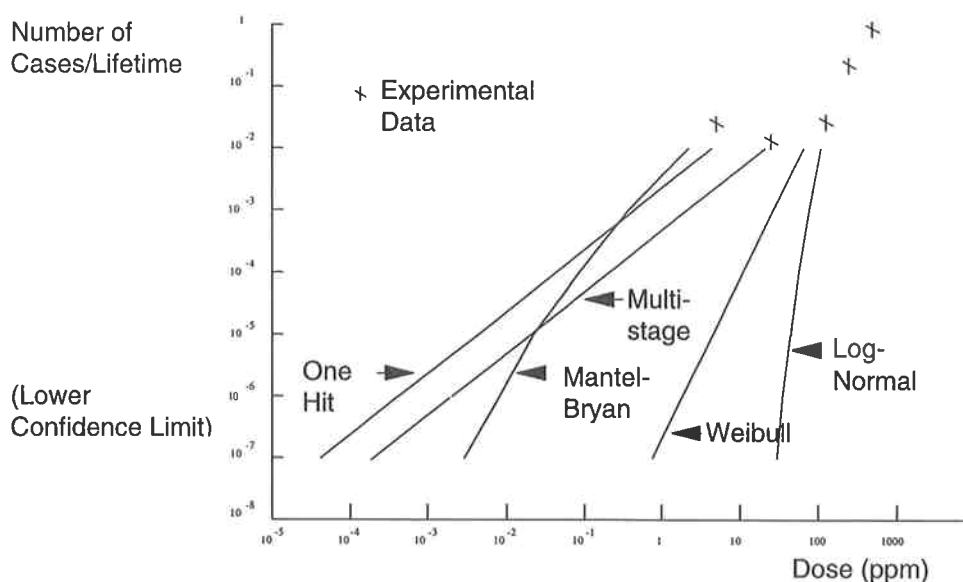
4.1 ESTIMATION OF RISK FROM SEVERAL MODELS

In a practical situation several models might adequately fit any particular set of experimental data, leaving little evidence on which to reject a particular model on the grounds of goodness of fit. It might be expected in these circumstances that the consequent estimates of risk should be reasonably consistent between the models. However, the slope of the extrapolated linear dose portion of the dose response curve differs for each of the models used. As a result the estimated RSD can vary by several orders of magnitude for the same experimental data. In general it is possible to rank the risk assessments either by RSD or potency estimated by the different models (Figure 4). They generally follow 1-Hit > Linear > Linearised Multi-stage > Weibull > Logit >

² The program used was TOX-RISK3 (Clement International Corporation, USA). Using the multi-stage model where q_1 (MLE) is zero, TOX-RISK makes an estimate of the MLE (examples 1, 3-6, 18, in Table 4).

Probit in their risk estimates for linear and sub-linear response curves (Johannsen, 1990). This relationship holds good for a wide range of chemicals, and is elegantly illustrated by Krewski and Van Ryzin (1981) and Munro and Krewski (1981) for saccharin, 2-acetylamino-fluorene (2-AAF) and aflatoxin. An example is shown in Figure 4 for ethylene thiourea (Cothurn, 1986). Often the multi-stage model is close to the linear model, and to the model free approach (Krewski *et al*, 1991).

Figure 4: Estimation of Cancer Risk by Different Mathematical Models from a single set of Experimental Data calculated on Tox-Risk 3 for Ethylenethiourea (redrawn from Cothurn, 1986)



4.2 THE APPLICATION OF THE MAXIMUM LIKELIHOOD ESTIMATE, MULTI-STAGE AND THE WEIBULL MODELS

The calculated q_1 (based on the MLE) and q_1^* (based on the 95% UCL) are key parameters in the estimation of risk, but there are certain limitations associated with the use of either which are discussed below.

4.2.1 Instability and Sensitivity of the Maximum Likelihood Estimation (MLE)

In theory the MLE represents the best estimate of risk. However, the MLE is unstable to minor fluctuations in the data. In particular, the additional incidence of a single tumour at another dose level in a bioassay can change the value of MLE from zero to a positive number with consequent changes in the RSD of 3-4 orders of magnitude (examples 1 and 2, 6 and 7, Table 2, see page 34).

4.2.2 Insensitivity of the Multi-stage Model

In the range of data exemplified in Table 2, the value of q_1 varies by over 5 orders of magnitude (2×10^{-9} to 6×10^{-3}) whereas the estimates for q_1^* from the same data set vary by less than one order of magnitude (3×10^{-3} to 1×10^{-2} in the multi-stage model). Given the large differences in the data sets used, this relatively small range is indicative of the insensitivity of q_1^* .

4.2.3 Lack of Discrimination in the Multi-stage Model

The tumour incidences exemplified in Table 2 include data which would be generally regarded as positive and sets which would be termed negative or equivocal (examples 3, 4 and 5) as well as examples with a strong positive trend in the absence of statistical significance between the high dose and control (example 20). The calculated value of q_1^* is the same for the positive or equivocal examples (see examples 1, 4 and 5) where in the latter examples there is only a small increment of tumour incidence at the top dose over a significant base line value. A value of q_1^* is obtained even when q_1 is zero. Irrespective of the data analysed the RSD (based on q_1^*) indicates lack of discrimination between the very different profiles exemplified. A wider (and less conservative) range of estimates is obtained by using q_1 or MLE.

4.2.4 Use of an Alternative Model (Weibull)

The problems encountered are not confined to the multi-stage model.

Using Weibull, the model did not fit two data examples (1 and 18, Table 2). A wide range (2×10^{-9} to 1×10^{-2}) of MLE was obtained for the remaining examples and the equivalent RSD relating to a 1×10^{-6} risk also covered a wide range. However, the RSD based on the 95% UCL, MLE varied by less than threefold (2×10^{-4} to 7×10^{-5}). As with the multi-stage model there was no clear discrimination between the negative or equivocal data and clearly positive data. Also, like the multi-stage model, the occurrence of one extra tumour (compare examples 6 and 7) can make 3-4 orders of magnitude of difference on the MLE and its resultant RSD. Examples 2, 3, 8, 9, 10 and 16 gave nearly identical estimates of risk for both the multi-stage and Weibull models. This is due to a constraint in the

model used preventing the Weibull dose-response curve becoming supra-linear at low dose (below the experimental range). At this point the program becomes similar to the linearised multi-stage model.

4.3 VARIATION OF DOSE RANGE AND RESPONSE CURVE

Risk estimates resulting from mathematical models should be dependent on the slope of the dose response curve and hence should reflect the differences between similar sets of incidence data with different dose spacing.

To test the sensitivity of the linearised multi-stage model, the same range of hypothetical tumour data was analysed, assuming it was derived from two separate dose levels (Table 3, see page 36). Dose range I (also used in the model in Table 2) had a 10-fold spacing between doses. This was compared to dose range II, with the top dose retained at the same nominal value, but the intermediate dose at $\frac{1}{2}$ the top dose level and the low dose at $\frac{1}{10}$ of the top dose level. This second range coincides more closely with the generalised recommendations for dose setting of IARC and reflects common practice in the bioassays conducted under the National Toxicology Programme (NTP) in the USA. The effective differences are: 1) a ten-fold increase in the low dose and 2) a higher intermediate dose. The results of the comparison are shown in Table 3. Dose range II showed a zero MLE for more of the tumour ranges exemplified and hence the RSD estimates based on this are lower. However, there is very little difference in the range of RSD (95% LCL) based on the q_1^* .

Clearly, there is only a small level of discrimination between the two data sets. With each individual example, the variation between the sets is generally 3-fold or less, a difference which is not significant in most estimates of risk.

4.4 EFFECT OF TOP DOSE

In an ideal model the estimate of risk would be less dependent on the absolute value of top dose than on the slope of the dose response curve. It has been shown in Table 3 that q_1^* and its RSD vary little with respect to a given tumour incidence for very different dose spacing. The high dose value also influences the estimates of risk. This is illustrated in Table 4 where for the same hypothetical tumour profile two dose ranges have been compared for the risk estimates which they generate. Dose range I corresponds to the data set used in Tables 2 and 3, whereas dose range II has a top dose an order of magnitude greater but retains the same intermediate and low dose. It may be seen that there is a decrease in the estimate of potency and an increase of the RSD by approximately 10-fold for all data

used. Thus in this model, other things being equal, risk estimates show a direct proportionality to the high dose selected in the bioassay.

4.5 RELATIONSHIP OF MAXIMUM TOLERATED DOSE AND CARCINOGENIC POTENCY

The proportionality between the high dose selected and the risk estimate exemplified in 4.4 is further complicated by the use of the Maximum Tolerated Dose (MTD) in most conventional (as carried out for regulatory purposes) bioassays. There is current debate on the exact method by which an MTD should be defined and a range of regulatory guidelines exists, some of which are mutually contradictory (see ECETOC, 1996). However, all agree that the dose should be at or close to the highest dose which may be given without altering the normal lifespan due to effects other than cancer. It is recognised that in a significant number of cases the mechanism of distribution and disposition of the chemical may differ between these high doses used and the lower doses more likely to be encountered in practice (NTP, 1984). In the extensive NCI/NTP database almost half the animal carcinogens would not have been identified if the MTD had not been used (Haseman, 1985). Nevertheless, the use of the MTD is defended on the grounds that the conventional bioassay is designed to identify the intrinsic hazard of the substance and not designed to estimate risk.

Good correlation has been demonstrated between MLE and UCL for chemicals identified as carcinogenic in rats and mice using both a one hit and a linearised multi-stage model (Crouch and Wilson, 1979; Crouch, 1983, Crouch *et al*, 1987; Gaylor and Chen, 1986; Chen and Gaylor, 1987; Gold *et al*, 1987; Rieth and Starr, 1989 a,b). High correlation between potency estimates for rats and mice, which have been shown to be dependent on high correlation between the MTDs in the same studies, have been used as a justification for quantitative extrapolation to human risk from animal studies. However model-design constraints restrict the value of the MLE within a very narrow range (a single order of magnitude using a linearised multi-stage model) of the inverse of the MTD. The high correlation between the MLE and the inverse of the MTD shown with both the one-hit and the linearised multi-stage model is therefore not surprising (Bernstein *et al*, 1985; Rieth and Starr, 1989a,b). As it is possible to fit a multi-stage model to data from a chemical which is not carcinogenic (see section 4.2.3) this relationship places a severe restraint on the range of values which can emerge from a risk assessment.

Tumours seen only at the MTD will greatly influence quantitative risk assessment (see Section 4.4), although their occurrence may be secondary to other manifestations of toxicity (ECETOC, 1996). The relevance of this dose for risk assessment must be questioned in view of the profound influence on the overall risk estimate and the confounding factors which may influence tumour incidence at such high doses.

4.6 LIMITATIONS OF THE DATA SETS USED

Most carcinogens detected in a bioassay have an increased tumour incidence at the high dose only. The exact proportion in this category depends on the statistical constraints applied, but Gold *et al* (1993) showed that up to 50% of chemicals showing carcinogenic activity in the NTP bioassays were inactive at the mid dose (set for these studies at half the top dose). For these chemicals only a single data point is available for modelling. Even where there are two data points available (the case with bioassays on most established human carcinogens - Apostalou 1990) these are insufficient to model a biological process other than as a straight line. Few if any chemicals in general commerce show an oncogenic response at more than two data points because such a dose relationship is rare for any chemical and because a NOAEL is a desired outcome for a regulatory study.

For a scientifically robust mathematical model at least three data points within the statistically reliable portion of the dose-response curve are needed. In theory some information could be derived from doses defined as NOAEL by assigning an upper bound to the possible response and using this value in the model. However the relatively small groups of the rodent bioassay mean that the calculated value would distort rather than enhance the model.

A standard carcinogenicity bioassay has the sensitivity to detect an increased tumour incidence of about 10%. The equations used extrapolate from these data to dose levels predicted to produce tumour incidence of the order of 1×10^{-6} (0.000,1%). Modelled extrapolations over even a single order of magnitude are tenuous, extrapolations over many orders of magnitude are not meaningful. The nature of the bioassay limits the range of available data to at best three orders of magnitude and (if regulatory guidelines are followed to the letter) more usually to only one order of magnitude.

5. DISCUSSION

5.1 DESIGN OF CONVENTIONAL BIOASSAYS

Conventional carcinogenicity bioassays are designed to identify the intrinsic hazardous (carcinogenic) properties of a chemical and not for quantitative estimation of risk. The standard experimental design of a control and 2 to 3 dose groups with 50 animals per group is a compromise between a design for hazard identification and for dose-response estimation. In such a study with 50 animals per group per sex an incidence of approximately 10% (against a zero control incidence) is needed to attain statistical significance. It is difficult to see how this design could sensibly be improved to accommodate the needs of risk assessment. Even the "mega-mouse study", where more than 24,000 animals were exposed to the genotoxic carcinogen 2-AAF, failed to detect an increase in the tumour incidence of 1 % and less (ED01, 1979).

Conventional bioassays use the MTD in order to maximise the sensitivity of detection. However, tumours seen at the MTD may result from a variety of mechanisms (such as cellular toxicity and consequent cellular replacement) which may either enhance the tumour rate or cause tumour production by acting as a promotion step. Chronic toxicity is known to be a promoter of carcinogenesis in animals in a variety of forms (ECETOC, 1996). The effects of toxicity and cellular proliferation are especially critical for non-genotoxic carcinogens. In the extensive NCI/NTP database almost half of the animal carcinogens would not have been identified if the MTD had not been used (Haseman, 1985). This suggests there are at least some chemicals where there is a threshold dose. In 78% of the studies with statistically-significant positives only at the MTD, there was a non-significant increase in the same tumour type at the next lower dose which is usually only half of the MTD (Hoel *et al*, 1988).

The use of the MTD to increase sensitivity may have validity in hazard detection. When the concept is extended to be used in risk assessment, then, at least with current mathematical models, additional complications are introduced. It has been demonstrated in Section 4 that the ultimate risk assessment is more dependent on the highest dose selected than on the shape of the dose response curve. This is perhaps not surprising given the study design where there are generally only three test groups, one of which is designed to be a no-effect level, (a feature which severely limits the value of mathematical modelling as at least three data points are needed for a robust model). Animals dosed at the MTD may show atypical metabolism and/or associated chronic toxicity which may distort the tumour profile from that expected in practice. Extrapolation from this dose without some correction for the dose response rate is therefore liable to overestimate both the number of chemicals carcinogenic at less extreme doses and also the associated risk.

The examples described in Section 4 of this document illustrate a number of issues. The upper confidence limit (UCL) may not be the most appropriate term for risk estimation as it can differ from the MLE by several orders of magnitude. Further the UCL is linear across a wide dose range and therefore does not convey all of the dose response information. This is particularly important from studies where the MTD concept is applied (see above). However, the UCL is relatively insensitive to small changes in tumour incidence while the MLE can be extremely sensitive to such changes (Table 2) which are of doubtful biological significance. With negative or equivocal data, where the response or trend does not attain statistical significance, the MLE is much reduced, whereas the UCL is not (Johannsen, 1990; Rieth and Starr, 1989a) i.e. the UCL does not discriminate between positive and negative data.

5.2 USE OF MATHEMATICAL MODELS

An accurate estimation of the probability of developing cancer when exposed to a specific chemical carcinogen would be of great benefit. However, it is apparent that the many uncertainties surrounding the current use of mathematical models severely limits the value of the calculated estimates to a point where their worth must be questioned. The reasons for these uncertainties are outlined in previous sections.

All mathematical models essentially assume linearity at low dose as a feature of their calculation. This assumption will overestimate risk if the true response below the experimental range of the response is sublinear. Zeise *et al* (1987) support this assumption on the basis of the dose response relationship for the formation of DNA adducts which is linear or very nearly linear over a wide range of doses including those which may be relevant for human exposure. Examples include benzo(a)pyrene in stomach and aflatoxin in liver. Care is required in the interpretation of such data since there may be non-linear dose-response relationships hidden in the overall linear observation of adduct formation and DNA adduct formation may not be the sole mechanism of carcinogenesis. Nevertheless these observations suggest that at least for some compounds linearity at low doses may be a reasonable approximation.

A large bioassay on 4080 rats reported by Peto *et al* (1991 a,b) on N-nitrosodiethylamine and N-nitrosodimethylamine provides some support for this assumption of linear response at low doses. Nevertheless this remains a contentious issue in the application of current mathematical models and may be a source of serious error in risk estimates, particularly when positive results are confined to the high dose.

By contrast Bailar *et al* (1988) have argued that for 308 chemicals tested by NCI/NTP, the one-hit model underestimates lifetime cancer risk in the observable range of the bioassay for a significant

fraction of the chemicals, suggesting that the low dose responses were supra-linear. However, Hoel and Portier (1992) in a more comprehensive analysis of the NCI/NTP database indicated a greater tendency towards sub-linearity, indicating that a linear assumption can overestimate risk.

More recently Sielken *et al* (1995) have suggested that for many compounds which are capable of inducing specific metabolic pathways, low dose levels may produce a hormetic effect i.e. the risk is actually reduced at low doses compared to the control. This theory is an extension of the "invaders" and "defenders" theory of Sielken (1987) and provides an alternative mathematical model for low dose extrapolations. The phenomenon had previously been suggested to hold good for many situations including ionising radiation (Wolff, 1989) and there is some experimental evidence in protozoa (Planel *et al*, 1987) to support this position. The model would impose an effective threshold even for genotoxic carcinogens.

Thus, there is no firm evidence on which to base the shape of the dose response curve below the observable range in animal studies. In the absence of data to the contrary, it has been considered reasonable to assume a linear response at low doses

A further potential deficiency in mathematical modelling is that the models are not able to take into account dose rate for a similar cumulative dose. If the concept of "invaders" and "defenders" of Sielken (1987) has any validity, then this may be of considerable significance as the use of mathematical models may overestimate risk from low level exposures e.g. from the diet but may significantly underestimate the consequences of episodic exposures such as may be encountered in an occupational setting.

5.3 CURRENT APPROACHES TO RISK ASSESSMENT

In most European countries scientific experts make a judgement on a case by case basis, considering all the evidence available, to derive an exposure level unlikely to lead to an increased cancer incidence in man. This can lead to over conservative decisions based on relatively inadequate information. Regulations then seek to reduce exposures to levels as low as reasonably practicable or eliminate it completely. This approach does not allow the estimation of an "acceptable risk" and can therefore be overzealous in regulation. Risk communication may also be difficult as the specific risk is not definable in recognisable terms. However, this system does provide an incentive to accumulate additional data as these improve the accuracy of the risk assessment.

The other approach commonly used, e.g. in the US and The Netherlands, seeks to estimate a lower limit for the dose associated with a specified increased lifetime risk of inducing cancer by using mathematical models to calculate potency figures. The limitations of this process are formally

recognised and the figure obtained is officially stated as "a plausible upper limit to the risk" (US-EPA, 1986). Nevertheless the single figure risk estimate gives a spurious sense of accuracy and the qualifications tend to be forgotten. However, the process may facilitate risk communication as the risk can be quantified in recognisable terms. For many chemicals the calculated extreme upper limit on risk presents no practical problem.

A report by the US Government Office of Management and Budget (OMB) in 1990 (CRA, 1991) criticised the quantitative risk assessment procedures used by the EPA and FDA as follows:

- the continued reliance on conservative (worst-case) assumptions by the Federal Agencies distorts the risk assessment giving estimates that exceed likely risks by several orders of magnitude;
- the generation and use of conservative biases provide a substantial margin of safety (and of uncertainty?), which is in the realm of risk management rather than risk assessment;
- the conservatism in the current risk assessment process distorts the regulatory priorities such that some chemicals presenting a possibly trivial cancer risk receive undue attention at the expense of other chemicals presenting more substantial threats to life and health.

On the other hand the American Industrial Health Council (AIHC) concluded that mathematical modelling was a potentially useful tool in risk assessment although it did not support the use of the linearised multi-stage model. The Council recommended a comprehensive review and update of the risk assessment process (AIHC, 1990 and 1993).

As a consequence, the US-EPA has embarked on a comprehensive review of the current risk assessment procedures and has proposed a process which takes into account all the available scientific evidence in an integrated manner as well as (and sometimes instead of) embarking on a quantitative risk assessment using a mathematical model.

5.4 IMPROVEMENTS IN RISK ASSESSMENT

The limitations of the risk assessment process (however practised) are widely recognised and improvements are constantly sought in all phases of the process. Perhaps the most significant area in which change is currently made is that of hazard evaluation and classification. Increasingly, a weight of evidence approach is being taken concerning classification. In particular this approach recognises the distinction between non-genotoxic carcinogens and genotoxic carcinogens. The characteristics of a suspect carcinogen can provide a number of clues concerning its mode of action

(ECETOC, 1983, 1990). Evidence for genotoxicity of a suspected carcinogen is obviously a key feature for classification although the distinction is not always clear cut. Thus, there are some animal carcinogens which are genotoxic *in vitro* and which have either equivocal or no activity *in vivo* e.g. methylene chloride, or which also require a second stimulus such as irritation e.g. formaldehyde to express their carcinogenicity. In these circumstances investigations of a non-genotoxic mechanism of carcinogenic action are usually made.

In the case of non-genotoxic carcinogens, it appears that one or more defined toxicological events precede the development of tumours thus presenting a threshold level of effect. Cohen and Ellwein (1990,1992) have advanced the concept of cell proliferation as a step in the carcinogenic process for some non-genotoxins. They also indicate that genotoxic and non-genotoxic (cell proliferation) mechanisms may act in parallel at higher doses for some compounds, for example for 2-AAF and bladder cancer in rats. The other major category of non-genotoxic mechanisms covers those which are receptor mediated. The non-genotoxic carcinogen TCDD acting via a cell receptor mechanism is one of the most potent animal carcinogens known. There also are self-evidently threshold doses as can be demonstrated by the magnitude of the signal required to elicit a response. However, non-genotoxicity is not necessarily equivalent to low potency.

It is currently assumed that genotoxic carcinogens do not exhibit a threshold dose. This has been challenged (see 5.1 and Sielken, 1987; Sielken *et al*, 1995) and Weisburger (1990) has claimed the existence of a threshold dose for the genotoxic carcinogens 2-AAF and benzo(a)pyrene. Nevertheless, at present the most widely held opinion is that, from the limited knowledge of the mechanism of action of genotoxic carcinogens, a threshold dose cannot be defined. As this is a key feature of risk assessment, further investigation in this area would clearly be productive.

In the dose-response and exposure-assessment phases, concepts such as exposure, route, dose, duration and concentration are considered both qualitatively and quantitatively. More complex sets of data may be considered in this phase apart from dose and route. These include effects on different species, target organ toxicity, metabolism and toxicokinetics, mechanistic studies and effects upon man. More refined approaches to the estimation of exposure such as using a distributional approach will lead to better-defined risk assessment. The importance of dose rate as opposed to cumulative dose may also need to be considered for future risk assessment.

5.5 ALTERNATIVE APPROACHES TO BASIC MATHEMATICAL MODELLING IN RISK ASSESSMENT FOR GENOTOXIC CARCINOGENS

This section deals with risk assessment of genotoxic carcinogens. It is generally accepted that the risks from exposure to non-genotoxic carcinogens may be assessed by using a safety factor (assessment factor) approach.

5.5.1 Biologically Motivated Models

Currently, the biologically-based models such as the Moolgavkar-Venzon-Knudson (MVK) models are not suitable for routine use as they require much more specific data than are currently obtained in bioassays. Studies such as those by Peto *et al* (1991 a,b) on N-nitrosodiethylamine and N-nitrosodimethylamine show the study size and intensity of observation required for such data to be obtained. Even in these circumstances it is not possible to extrapolate with confidence to low doses. Currently these models can be used only on a case-by-case basis where suitable data exist. Experience may allow these models to be simplified and developed further.

5.5.2 Assessment Factor Approach

The use of a mathematical model for extrapolating below the experimental dose range is in practice equivalent to applying a safety factor multiplied by a large and ill-defined error factor.

The simplest alternative would be to apply an assessment factor to the measured NOAEL or to a LOAEL when this was the lowest level tested. The assessment factor used should be derived on a case-by-case basis, considering the nature of the chemical and the population being exposed and may be higher or lower than the conventional 100-fold value. A value range from 10 to 5,000 has been suggested (Renwick, 1991; Kuiper-Goodman, 1991; Lu and Sielken, 1991).

The relationship between MTD and potency has led to the observation that dividing the MTD by a factor of 380,000 will approximate to a dose relating to a risk of 1×10^{-6} using a linearised multi-stage model. This approach is effectively a restatement of the limitation of the model (see section 4.5) and suffers from the problems previously described.

5.5.3 Bench Mark Dose

The concept of a bench mark dose (BMD) has been investigated in other areas e.g. developmental toxicology (Gaylor, 1989; US EPA, 1991c). The basis of this approach, is a mathematical model fitted to the experimental data within the observable range to estimate a dose corresponding to a

defined level of effect, such as 1, 5 or 10% increase in the incidence of a specified effect (ED01, ED05, or ED10). As a 10% increase is about the smallest change that is statistically significant in a standard cancer bioassay, the ED10 is appropriate for cancer data. Using a BMD that is within (or at worst very close to) the observable range of the experiment, reduces the problems associated with dose extrapolation. Estimates of the benchmark dose (e.g. the ED10, or its lower confidence limit, LED10) would reflect the doses at which changes in tumour incidence occurred, and are quite insensitive to the mathematical model used. It must be recognised that the benchmark model will suffer from the same data limitations as the other models in current use. Thus under normal circumstances only one, or at best two, doses are available for model development. This may be the reason for the various models producing similar results rather than an inherent robust property of the BMD and the use of this model may still lead to over-conservative evaluations. Nevertheless, where appropriate data exist the benchmark dose may find use in risk assessment as an apparently robust measure of tumour potency and could be combined with appropriate assessment factors to set acceptable levels for human exposure. The model would encourage the generation of additional data points to improve extrapolation and may therefore remain only applicable to shorter term studies with smaller group size.

5.5.4 Concept of "Threshold of Regulation"

Krewski *et al* (1991) have reviewed a concept of a "threshold of regulation" for chemical carcinogens. Based on data obtained from the carcinogen potency data base (CPDB) for 585 experiments (Gold *et al*, 1989), the dose corresponding to a 1×10^{-6} increase in risk was roughly log-normally distributed around a median of 70-90 ng/kg/d. Exposure to dose levels greater than this range would be considered unacceptable. The dose was estimated by linear extrapolation from the TD_{50} and was within a factor of 5-10 of the figure obtained from the linearised multi-stage model. The TD_{50} 's are related to the MTD (see Section 4.5) so that the Threshold of Regulation concept is effectively a regulation on the basis of frank toxicity rather than carcinogenicity.

Such an approach has no apparent biological basis and would require much deeper consideration of biological, analytical and mathematical issues and a much wider data base for validation. Further investigation into measures of potency for carcinogens may throw further light onto this area.

6. RECOMMENDATIONS

With rare exceptions mathematical models such as the linearised multi-stage cannot be recommended for carcinogen risk assessment. Other types of mathematical models, incorporating more sophisticated approaches involve PB-PK and/or biologically-based modelling. These may be useful on a case-by-case basis to define more accurately the limits for the cancer risk of an individual substance at the low-dose levels to which human may be exposed.

Based on the above considerations it is recommended that:

1. Initial classification of a carcinogen should be followed by a thorough investigation of data using both qualitative and quantitative approaches as exemplified by the UK DOH (1991) Carcinogenicity guidelines which state:

"The Committee on Carcinogenicity evaluates data on a case-by-case basis, taking into account the weight of all available evidence. The range of data considered may differ with the circumstances and it is not possible to provide a universally applicable list which will be needed for a carcinogenicity risk assessment. When a plausible risk of cancer is confirmed at this review then an assessment factor should be applied to the estimated Virtually Safe Dose (VSD) to define a limit value for human exposure."

2. Variable assessment factors are used on a case-by-case basis and the application of assessment factors in carcinogenicity risk assessment should be thoroughly reviewed to assess their potential for wider use than is currently practised.
3. Alternative approaches are considered which avoid complex mathematical extrapolation beyond the experimental range. Those currently under development including linear model-free and bench-mark dose models require further evaluation to determine their potential (if any) in carcinogen risk assessment.

In using risk assessment procedures it should be remembered that, whilst not addressed in this document, exposure assessment (Stage 5 in Fig. 1) presents another uncertainty factor which may increase the range of estimates when assessing human cancer risk.

Table 2: Comparison of Risk Estimates using Two Models

Example No.	Tumour Incidence				Multi-stage Model				Weibull Model			
					a	b	c	d	a	b	c	d
	con	low	mid	high	q ₁ (MLE)	RSD (MLE) mg/kg/d	q ₁ [*] (95%UCL)	RSD (95% LCL) mg/kg/d	q ₁ [*] (MLE)	RSD(MLE) mg/kg/d	q ₁ [*] 95%UCL	RSD (95% LCL) mg/kg/d
1	0	0	0	5	0 (1x10 ⁻⁵)	0 (1)	3x10 ⁻³	3x10 ⁻⁴	0	Model does not fit	Model does not fit	Model does not fit
2	0	0	1	5	2x10 ⁻³	4x10 ⁻⁴	4x10 ⁻³	2x10 ⁻²	2x10 ⁻³	4x10 ⁻⁴	4x10 ⁻³	2x10 ⁻⁴
3	5	5	5	5	0 (2x10 ⁻⁹)	0 (5x10 ⁻¹²)	2x10 ⁻³	5x10 ⁻⁴	2x10 ⁻⁹	5x10 ⁻²	2x10 ⁻³	5x10 ⁻⁴
4	5	5	5	6	0 (6x10 ⁻⁷)	0 (2)	3x10 ⁻³	4x10 ⁻⁴	1x10 ⁻⁷	8	3x10 ⁻³	4x10 ⁻⁴
5	5	5	5	7	0 (7x10 ⁻⁷)	1	3x10 ⁻³	3x10 ⁻⁴	2x10 ⁻⁷	6	3x10 ⁻³	3x10 ⁻⁴
6	0	0	0	10	0 (1x10 ⁻⁶)	0 (8x10 ⁻¹)	5x10 ⁻³	2x10 ⁻⁴	5x10 ⁻⁸	2x10 ⁻¹	5x10 ⁻³	2x10 ⁻⁴
7	0	0	1	10	4x10 ⁻³	3x10 ⁻⁴	7x10 ⁻³	2x10 ⁻⁴	1x10 ⁻³	8x10 ⁻⁴	7x10 ⁻³	2x10 ⁻⁴
8	0	0	2	10	5x10 ⁻³	2x10 ⁻⁴	7x10 ⁻³	1x10 ⁻⁴	5x10 ⁻³	2x10 ⁻⁴	7x10 ⁻³	1x10 ⁻⁴
9	0	0	3	10	5x10 ⁻³	2x10 ⁻⁴	8x10 ⁻³	1x10 ⁻⁴	5x10 ⁻³	2x10 ⁻⁴	8x10 ⁻³	1x10 ⁻⁴
10	0	0	4	10	6x10 ⁻³	2x10 ⁻⁴	8x10 ⁻³	1x10 ⁻⁴	6x10 ⁻³	2x10 ⁻⁴	8x10 ⁻³	1x10 ⁻⁴

Models used hypothetical tumour incidence. Observed incidence (group size 50) is from control (left) to high dose (right) a: q₁ or maximum likelihood estimate (MLE), (where q₁ is zero the Tox-Risk estimate of MLE is given in parenthesis). b: Risk specific dose (RSD) given for 1 X 10⁻⁶ using q₁ or MLE. c: q₁^{*}, or Unit Potency Factor, (95% UCL on q₁) d: Risk Specific Dose (RSD) given for 1 X 10⁻⁶ risk using q₁^{*} (i.e. 95% LCL on risk).

Table 2 (cont.): Comparison of Risk Estimates using Two Models

Example No.	Tumour Incidence				Multi-stage Model				Weibull Model			
	con	low	mid	high	a q _i (MLE)	b RSD (MLE) mg/kg/d	c q _i * (95%UCL)	d RSD (95% LCL) mg/kg/d	a q _i (MLE)	b RSD(MLE) mg/kg/d	c q _i 95%UCL	d RSD (95% LCL) mg/kg/d
11	0	1	3	10	5x10 ⁻³	2x10 ⁻⁴	8x10 ⁻³	1x10 ⁻⁴	5x10 ⁻³	2x10 ⁻⁴	8x10 ⁻⁴	1x10 ⁻⁴
12	1	1	3	10	4x10 ⁻³	2x10 ⁻⁴	7x10 ⁻³	1x10 ⁻⁴	4x10 ⁻³	2x10 ⁻⁴	7x10 ⁻³	1x10 ⁻⁴
13	1	2	3	10	4x10 ⁻³	3x10 ⁻⁴	7x10 ⁻³	2x10 ⁻⁴	4x10 ⁻³	3x10 ⁻⁴	7x10 ⁻³	2x10 ⁻⁴
14	1	2	5	10	4x10 ⁻³	3x10 ⁻³	7x10 ⁻³	1x10 ⁻⁴	4x10 ⁻³	2x10 ⁻⁴	7x10 ⁻³	1x10 ⁻⁴
15	1	3	5	10	4x10 ⁻³	3x10 ⁻⁴	7x10 ⁻³	2x10 ⁻⁴	5x10 ⁻³	2x10 ⁻⁴	7x10 ⁻³	1x10 ⁻⁴
16	0	0	7	14	9x10 ⁻³	1x10 ⁻⁴	1x10 ⁻²	8x10 ⁻⁵	9x10 ⁻³	1x10 ⁻⁴	1x10 ⁻²	8x10 ⁻⁵
17	0	6	8	15	6x10 ⁻³	2x10 ⁻⁴	1x10 ⁻²	1x10 ⁻⁴	1x10 ⁻²	1x10 ⁻⁴	1x10 ⁻²	8x10 ⁻⁵
18	5	5	5	15	0 (1x10 ⁻⁶)	0 (8x10 ⁻¹)	9x10 ⁻³	1x10 ⁻⁴	Model does not fit			
19	5	5	12	19	8x10 ⁻³	1x10 ⁻⁴	1x10 ⁻²	8x10 ⁻⁵	8x10 ⁻³	1x10 ⁻⁴	1x10 ⁻²	8x10 ⁻⁵
20	5	11	13	20	7x10 ⁻³	2x10 ⁻⁴	1x10 ⁻²	9x10 ⁻⁵	9x10 ⁻³	1x10 ⁻⁴	1x10 ⁻²	7x10 ⁻⁵

Models used hypothetical tumour incidence. Observed incidence (group size 50) is from control (left) to high dose (right) a: q_i or maximum likelihood estimate (MLE), (where q_i is zero the Tox-Risk estimate of MLE is given in parenthesis). b: Risk specific dose (RSD) given for 1 X 10⁻⁶ using q_i or MLE. c: q_i*, or Unit Potency Factor, (95% UCL on q_i) d: Risk Specific Dose (RSD) given for 1 X 10⁻⁶ risk using q_i* (i.e. 95% LCL on risk).

Table 3: Effect Of Low Dose Value On Risk Estimates -

Example No.	Tumour Incidence				Dose Range I (0, 0.5, 5 & 50mg/kg/d)				Dose Range II (0, 5, 25 & 50 mg/kg/d)			
	con	low	mid	high	a q _i (MLE)	b RSD (MLE) mg/kg/d	c q _i * (95%UCL)	d RSD (95% LCL) mg/kg/d	a q _i (MLE)	b RSD (MLE) mg/kg/d	c q _i * 95%UCL	d RSD (95% LCL) mg/kg/d
1	0	0	0	5	(1X10 ⁵)	0 (1)	3X10 ³	3X10 ⁴	0 (9X10 ⁻⁷)	0 (1)	1X10 ³	8X10 ⁻⁴
2	0	0	1	5	2X10 ³	4X10 ⁻⁴	4X10 ³	2X10 ⁴	0 (4X10 ⁻⁶)	2X10 ⁻¹	2X10 ³	4X10 ⁻⁴
3	5	5	5	5	0 (2X10 ⁻⁹)	0 (5X10 ⁻¹²)	2X10 ³	5X10 ⁴	0 (2X10 ⁻⁹)	0 (5X10 ⁻²)	2X10 ³	5X10 ⁻⁴
4	5	5	5	6	0 (6X10 ⁻⁷)	0 (2)	3X10 ³	4X10 ⁴	0 (6X10 ⁻⁷)	0 (2)	3X10 ³	4X10 ⁻⁴
5	5	5	5	7	0 (7X10 ⁻⁷)	0 1	3X10 ³	4X10 ⁴	0 (7X10 ⁻⁷)	0 (1)	3X10 ³	3X10 ⁻⁴
6	0	0	0	10	0 (1X10 ⁻⁶)	0 (8X10 ⁻¹)	5X10 ³	2X10 ⁴	0 (1X10 ⁻⁶)	0 (9X10 ⁻¹)	1x10 ³	9X10 ⁻⁴
7	0	0	1	10	4X10 ³	3X10 ⁻⁴	7X10 ³	2X10 ⁴	0 (1X10 ⁻⁶)	0 (8X10 ⁻¹)	2X10 ³	5x10 ⁻⁴
8	0	0	2	10	5x10 ³	2x10 ⁻⁴	7X10 ³	1X10 ⁴	0 (6X10 ⁻⁶)	0 (2X10 ⁻¹)	3X10 ³	3X10 ⁻⁴
9	0	0	3	10	5X10 ³	2X10 ⁻⁴	8X10 ³	1X10 ⁴	0 (1X10 ⁻⁵)	0 (1X10 ⁻¹)	4X10 ³	3X10 ⁻⁴
10	0	0	4	10	6X10 ³	2x10 ⁻⁴	8x10 ³	1x10 ⁴	1X10 ⁻⁵	1X10 ⁻¹	5X10 ³	2X10 ⁻⁴

Table 3 (cont.): Effect Of Low Dose Value On Risk Estimates -

Example No.	Tumour Incidence				Dose Range I (0, 0.5, 5 & 50mg/kg/d)				Dose Range II (0, 5, 25 & 50 mg/kg/d)			
	con	low	mid	high	a q _i (MLE)	b RSD (MLE) mg/kg/d	c q _i [*] (95%UCL)	d RSD (95% LCL) mg/kg/d	a q _i ¹ (MLE)	b RSD(MLE) mg/kg/d	c q _i ¹ 95%UCL	d RSD (95% LCL) mg/kg/d
11	0	1	3	10	5X10 ³	2x10 ⁻⁴	8X10 ⁻³	1X10 ⁻⁴	3X10 ⁻³	4X10 ⁻⁴	6X10 ⁻³	2X10 ⁻⁴
12	1	1	3	10	4X10 ³	2X10 ⁻⁴	7X10 ⁻³	1X10 ⁻⁴	0 (7X10 ⁻⁶)	0 (1X10 ⁻¹)	5X10 ⁻³	2X10 ⁻⁴
13	1	2	3	10	4X10 ³	3X10 ⁻⁴	7X10 ⁻³	2X10 ⁻⁴	9X10 ⁻⁴	1X10 ⁻³	5X10 ⁻³	2X10 ⁻⁴
14	1	2	5	10	4X10 ³	3X10 ⁻⁴	7X10 ⁻³	1X10 ⁻⁴	3X10 ⁻³	3X10 ⁻²	6X10 ⁻³	2X10 ⁻⁴
15	1	3	5	10	4X10 ³	3X10 ⁻⁴	7X10 ⁻³	2X10 ⁻⁴	4X10 ⁻³	3X10 ⁻⁴	6X10 ⁻³	2X10 ⁻⁴
16	0	0	7	14	9X10 ³	1X10 ⁻⁴	1X10 ⁻²	8X10 ⁻⁵	1X10 ⁻³	7X10 ⁻⁴	7X10 ⁻³	1X10 ⁻⁴
17	0	6	8	15	6X10 ³	2X10 ⁻⁴	1X10 ⁻²	1X10 ⁻⁴	7X10 ⁻³	1X10 ⁻⁴	1X10 ⁻²	9X10 ⁻⁵
18	5	5	5	15	0 (1X10 ⁻⁶)	0 (8X10 ⁻¹)	9X10 ⁻³	1X10 ⁻⁴	0 (1X10 ⁻⁶)	0 (8X10 ⁻¹)	5X10 ⁻³	2X10 ⁻⁴
19	5	5	12	19	8X10 ³	1X10 ⁻⁴	1X10 ⁻²	8X10 ⁻⁵	5X10 ⁻³	2X10 ⁻⁴	1X10 ⁻²	9X10 ⁻⁵
20	5	11	13	20	7X10 ³	2X10 ⁻⁴	1X10 ⁻²	9X10 ⁻⁵	7X10 ⁻³	1X10 ⁻⁴	1X10 ⁻²	9X10 ⁻⁵

Comparison of fitting the multi-stage model to a hypothetical set of tumour incidence (control on left to high dose on right) related to two different dose spacing methods employing the same top dose a.q_i or maximum likelihood estimate (MLE), (where q_i is zero Tox-Risk estimate of MLE is given in parenthesis). b; Risk specific dose (RSD) given for 1 x 10⁻⁶ risk. c: q_i^{*} or Unit Potency Factor (95% UCL on q_i). d: Risk Specific Dose (RSD) given for 1 x 10⁻⁶ risk based on q_i^{*} (i.e. 95% LCL on Risk).

Table 4: Effect Of High Dose Value On Risk Estimates

Example No.	Tumour Incidence				Dose Range I (0, 0.5, 5 & 50mg/kg/d)				Dose Range II (0, 5, 25 & 50 mg/kg/d)			
	con	low	mid	high	a q ₁ (MLE)	b RSD (MLE) mg/kg/d	c q ₁ [*] (95%UCL)	d RSD (95% LCL) mg/kg/d	a q ₁ [*] (MLE)	b RSD(MILE) mg/kg/d	c q ₁ [*] 95%UCL	d RSD (95% LCL) mg/kg/d
1	0	0	0	5	0 (1X10 ⁻⁶)	0 (1)	3X10 ⁻³	3X10 ⁻⁴	0 (9X10 ⁻⁶)	0 (10)	4X10 ⁻⁴	3X10 ⁻³
2	0	0	1	5	2X10 ⁻³	4X10 ⁻⁴	4X10 ⁻³	2X10 ⁻⁴	2X10 ⁻⁴	5X10 ⁻³	4X10 ⁻⁴	2X10 ⁻³
3	5	5	5	5	0 (2X10 ⁻⁹)	0 (5X10 ⁻¹²)	2X10 ⁻³	5X10 ⁻⁴	0 (2X10 ⁻¹⁰)	0 (5X10 ⁻⁹)	2X10 ⁻⁴	5X10 ⁻³
4	5	5	5	6	0 (6X10 ⁻⁷)	0 (2)	3X10 ⁻³	4X10 ⁻⁴	0 (6X10 ⁻⁶)	0 (20)	3X10 ⁻⁴	4X10 ⁻³
5	5	5	5	7	0 (7X10 ⁻⁷)	0 (1)	3X10 ⁻³	3X10 ⁻⁴	0 (7X10 ⁻⁶)	0 (20)	3X10 ⁻⁴	3X10 ⁻³
6	0	0	0	10	0 (1X10 ⁻⁶)	0 (8X10 ⁻¹)	5X10 ⁻³	2X10 ⁻⁴	0 (1X10 ⁻⁷)	0 (8)	7X10 ⁻⁴	1X10 ⁻³
7	0	0	1	10	4X10 ⁻³	3X10 ⁻⁴	7X10 ⁻³	2X10 ⁻⁴	4X10 ⁻⁴	2X10 ⁻³	7X10 ⁻⁴	1X10 ⁻³
8	0	0	2	10	5X10 ⁻³	2X10 ⁻⁴	7X10 ⁻³	1X10 ⁻⁴	4X10 ⁻⁴	2X10 ⁻³	7X10 ⁻⁴	1X10 ⁻³
9	0	0	3	10	5X10 ⁻³	2X10 ⁻⁴	8X10 ⁻³	1X10 ⁻⁴	4X10 ⁻⁴	2X10 ⁻³	7X10 ⁻⁴	1X10 ⁻³
10	0	0	4	10	6X10 ⁻³	2X10 ⁻⁴	8X10 ⁻³	1X10 ⁻⁴	6X10 ⁻⁴	2X10 ⁻³	8X10 ⁻⁴	1X10 ⁻³

Comparison of fitting the multi-stage model using a hypothetical tumour incidence (control on left, high dose on right) using logarithmic dose spacing differing by a factor of 10; a: q₁ or maximum likelihood estimate (MLE), (where q₁ is zero Tox-Risk estimate of MLE is given in parenthesis). b: Risk specific dose (RSD) given for 1 X 10⁻⁶ risk. c: q₁^{*} Unit Potency Factor (95% UCL on q₁), d: Risk Specific Dose (RSD) given for 1 X 10⁻⁶ risk based on q₁^{*} (95% LCL on risk).

Table 4 (cont.): Effect Of High Dose Value On Risk Estimates

Example No.	Tumour Incidence				Dose Range I (0, 0.5, 5 & 50mg/kg/d)				Dose Range II (0, 5, 25 & 50 mg/kg/d)			
	con	low	mid	high	a q ₁ (MLE)	b RSD (MLE) mg/kg/d	c q ₁ [*] (95%UCL)	d RSD (95% LCL) mg/kg/d	a q ₁ [†] (MLE)	b RSD(MLE) mg/kg/d	c q ₁ [†] 95%UCL	d RSD (95% LCL) mg/kg/d
11	0	1	3	10	5X10 ⁻³	2X10 ⁻⁴	8X10 ³	1X10 ⁻⁴	5X10 ⁻⁴	2X10 ⁻³	8X10 ⁻⁴	1X10 ⁻³
12	1	1	3	10	4X10 ⁻³	2X10 ⁻⁴	7X10 ³	1X10 ⁻⁴	4X10 ⁻⁴	2X10 ⁻³	7X10 ⁻⁴	1X10 ⁻³
13	1	2	3	10	4X10 ⁻³	3X10 ⁻⁴	7X10 ³	2X10 ⁻⁴	4X10 ⁻⁴	3X10 ⁻³	7X10 ⁻⁴	2X10 ⁻³
14	1	2	5	10	4X10 ⁻³	3X10 ⁻⁴	7X10 ³	1X10 ⁻⁴	4X10 ⁻⁴	3X10 ⁻³	7X10 ⁻⁴	1X10 ⁻³
15	1	3	5	10	4X10 ⁻³	3X10 ⁻⁴	7X10 ³	2X10 ⁻⁴	4X10 ⁻⁴	3X10 ⁻³	7X10 ⁻⁴	2X10 ⁻³
16	0	0	7	14	9X10 ⁻³	1X10 ⁻⁴	1X10 ²	8X10 ⁻⁵	9X10 ⁻⁴	1X10 ⁻³	1X10 ⁻³	8X10 ⁻⁴
17	0	6	8	15	6X10 ⁻³	2X10 ⁻⁴	1X10 ²	1X10 ⁻⁴	6X10 ⁻⁴	2X10 ⁻³	1X10 ⁻³	1X10 ⁻³
18	5	5	5	15	0 (1X10 ⁻⁶)	0 (8X10 ⁻¹)	9X10 ³	1X10 ⁻⁴	0 (1X10 ⁻⁷)	8	9X10 ⁻⁴	1X10 ⁻³
19	5	5	12	19	8X10 ⁻³	1X10 ⁻⁴	1X10 ²	8X10 ⁻⁵	8X10 ⁻⁴	1X10 ⁻³	1X10 ⁻³	8X10 ⁻⁴
20	5	11	13	20	7X10 ⁻³	2X10 ⁻⁴	1X10 ²	9X10 ⁻⁵	7X10 ⁻⁴	2X10 ⁻³	1X10 ⁻²	9X10 ⁻⁴

Comparison of fitting the multi-stage model using a hypothetical tumour incidence (control on left, high dose on right) using logarithmic dose spacing differing by a factor of 10; a: q₁ or maximum likelihood estimate (MLE), (where q₁ is zero Tox-Risk estimate of MLE is given in parenthesis). b: Risk specific dose (RSD) given for 1 X 10⁻⁶ risk. c: q₁^{*} Unit Potency Factor (95% UCL on q₁), d: Risk Specific Dose (RSD) given for 1 X 10⁻⁶ risk based on q₁^{*} (95% LCL on risk).

APPENDIX I: CLASSIFICATION SCHEMES FOR CARCINOGENIC SUBSTANCES

Scheme 1: Classification of Carcinogens as Used by IARC (1987)

Group	Definition	Description
1	Carcinogenic to humans	Sufficient evidence of carcinogenicity in humans
2a	Probably carcinogenic to humans	Limited evidence in humans; less than sufficient evidence in animals
2b	Possibly carcinogenic to humans	Limited evidence in humans; no sufficient evidence in animals or Inadequate/non-existent evidence in humans; sufficient evidence in animals
3	Not classifiable	Agents that are not categorised in any other group
4	Probably not carcinogenic	Evidence suggesting no carcinogenicity in humans or inadequate data; evidence suggesting no carcinogenicity in animals

Scheme 2: Classification of Carcinogens as Used by US-EPA (1986)

Category	Definition	Description
A	Known human carcinogen	Proven human carcinogenic substance
B1	Probable human carcinogen	Suspected human carcinogenic substance of potential relevance to humans
B2	Probable human carcinogen	Proved animal carcinogenic substance of potential relevance to humans
C	Possible human carcinogen	Suspected animal carcinogenic substance of potential relevance to humans
D	Not classifiable as to human carcinogenicity	Substances non-classifiable with regard to carcinogenicity
E	Evidence on non-carcinogenicity for humans	Negative evidence
None	No data

Scheme 3: Classification of Carcinogens as Used by EEC (1991)

Group	Definition	Description
1	Carcinogenic to man	Sufficient evidence of carcinogenicity in man
2	To be regarded as if carcinogenic to man	Sufficient evidence in animals sufficient additional evidence that the animal response is relevant for man
3	Concern owing to possible carcinogenic effects	Sufficient or limited evidence in animals evidence that the animal response may not be relevant for man
None	No classification	Absence of evidence for carcinogenicity evidence in animals with evidence that the animal response is not relevant for man

Scheme 4: Classification of Carcinogens as Proposed by ASHBY *et al* (1991)

Category	Definition	Description
1	Known human carcinogen	Sufficient evidence for human carcinogenicity
2	Probable human carcinogen	Limited (inadequate)evidence in humans; sufficient evidence in animals (strong evidence that animal response is relevant for humans)
3	Possible human carcinogen	Limited (inadequate) evidence in humans; limited evidence in animals (inadequate information concerning relevance of animal response for humans)
4	Equivocal evidence	Inadequate evidence in humans; limited evidence in animals; inadequate/no information concerning relevance of animal response for humans
5	Evidence inadequate for classification	Human and animal data inadequate or suggestive of non-carcinogenicity
6	Carcinogenic in animals; probably not a human carcinogen	Human evidence inadequate or suggestive of non-carcinogenicity; animal evidence sufficient/limited evidence that animal response is not relevant for humans
7	Carcinogenicity in animals	Human evidence inadequate or suggestive of non-carcinogenicity

BIBLIOGRAPHY

- AIHC, 1990. (American Industrial Health Council). An Approach to Comprehensive Methodology in Assessing Cancer Risk. Washington, D C.
- AIHC, 1993. (American Industrial Health Council). Improving and Updating Risk Assessment: Precepts for Legislative Proposals.
- Andersen ME, Clewell MJL and Reitz RH, 1987. Physiologically based pharmacokinetics and methylene chloride cancer risk assessment considerations. In: Human risk assessment, The Role of Animal Selection and Extrapolation. Ed. Roloff MV, Taylor and Francis, London, 129.
- Andersen ME, Krishnan K, Conolly RB and McClellan RO, 1992. Mechanistic toxicology research and biologically based modelling; partners for improving quantitative risk assessments. CIIT Bulletin 12(1), 1-8.
- Apostalou A, 1990. Relevance of maximum tolerated dose to human carcinogenic risk. Regul Toxicol Pharmacol 11, 68-73.
- Armitage P, 1982. The assessment of low dose carcinogenicity. Biometrics 38, 119-129.
- Armitage P and Doll R, 1961. Stochastic models for carcinogenesis in 'Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability. Eds. Lecam L. and Neyman, J. University of California Press., Berkeley, LA, 19-38.
- Ashby J, Doerrner NG, Flamm WG, Harris JE, Hughes DH, Johannsen FR, Lewis SC, Krivanek ND, McCarthy JF, Moolenaar RJ, Raabe GK, Reynolds RC, Smith JM, Stevens JT, Teta MJ and Wilson JD, 1991. A scheme for Classifying carcinogens. Regul Toxicol Pharmacol 12, 270-295.
- Atterwill CK, Jone C and Brown CG, 1992. Thyroid gland II-Mechanisms of species-dependent thyroid toxicity, hyperplasia and neoplasia induced by xenobiotics. In "Endocrine Toxicology" Eds. Atterwill C.K. and Flack J.D. pp137-182, Cambridge Univ. Press
- Bailar JC III, Crouch GAC, Shaikh R and Speigelman D, 1988. One-hit models of carcinogenesis: conservative or not? Risk Anal 8, 485-497.
- Berstein L, Gold LS, Ames BN, Pike MC and Hoel DG, 1985. Some tautologous aspects of the comparison of carcinogenic potency in rats and mice. Fund Appl Toxicol 5, 79-86.
- Bogen KT, 1988. Pharmacokinetics of regulations risk analysis: the case of trichloroethylene. Regul Toxicol Pharmacol 8, 447.
- Brown KB and Hoel DG, 1983. Multi-stage prediction of cancer in serially dosed animals with application to the ED01 study. Fund Appl Toxicol 3, 470-477.
- Chen JJ and Gaylor DW, 1987. Carcinogenic risk assessment: Comparison of estimated safe doses for rats and mice. Environ Health Perspect 72, 305-309.
- CIA, 1991. Hazard Assessment of New Chemicals, an introduction. Chemical Industries Association, London.
- Cohen SM and Ellwein LB, 1990. Cell Proliferation in Carcinogenesis. Science 249, 1007-1011.
- Cohen SM and Ellwein LB, 1992. Risk assessment based on high dose animal exposure experiments. Chem Res Toxicol 5, 742-748.
- Conolly RB and Andersen ME, 1991. Biologically Based Pharmacodynamic Models: Tools for toxicological research and risk assessment. Ann Rev Pharmacol Toxicol 31, 503-523.
- COSHH, 1988. Control of Hazardous Substances to Health (COSHH), Approved Code of Practice. HMSO, London.

- Crouch EAC, 1983. Uncertainties in interspecies extrapolations of carcinogenicity. *Environ Health Persp* 50, 321-327.
- Crouch EAC and Wilson R, 1979. Interspecies comparison of carcinogenic potency. *J Toxicol Environ Health* 5, 1095-1118.
- Crouch EHC, Wilson R and Zcise L, 1987. Tautology or not tautology. *J Toxicol Environ Health* 20, 1-10.
- Crump KS, Guess HA and Deal KL, 1977. Confidence intervals and tests of hypothesis concerning dose response relations inferred from animal carcinogenicity data. *Biometrics* 33, 437-451.
- Crump KS, 1984. An improved procedure for low-dose carcinogenic risk assessment from animal data. *J Environ Pathol Toxicol Oncol* 6, 339-348.
- D'Souza RW, Francis WR, Bruce RD and Andersen ME, 1987. Physiologically based pharmacokinetic model for ethylene dichloride and its application in risk assessment. In: *Pharmacokinetics in Risk Assessment. Drinking Water and Health*, National Academy & Press, Washington DC, 286.
- EEC, 1991. 91/325/EEC. Commission Directive of 1 March 1991 adapting to technical progress for the 12th time Council Directive 67/548/EEC on the approximate of the laws, regulations and administrative provisions to the classification, packaging and labelling of dangerous substances. *O.J. Eur. Comm.* 1991, L180 34, page L 180/59. Section 4.2.1.2.1.
- EEC, 1992. Council Directive 92/32/EEC of 30 April 1992 Amending for the Seventh Time Directive 67/548/EEC on the approximation of the laws.
- EEC, 1994. Commission Regulation (EC) No 1488/94 of 28 June 1994 laying down the principles for the assessment of risks to man and the environment of existing substances in accordance with Council Regulation (EEC) No 793/93.
- EEC, 1993. Council Regulation (EEC) No 793/93 of 23 March 1993 on the evaluation and control of the risks of existing substances.
- ECETOC, 1983. Risk Assessment of Occupational Chemical Carcinogens. Monograph N°3.
- ECETOC, 1986. Structure Activity Relationships in Toxicology and Ecotoxicology: An assessment. Monograph N°8.
- ECETOC, 1988. Methylene chloride (dichloromethane): Human risk assessment using experimental animal data. Green T, Proven WM, Dugard PH, Cook SK.
- ECETOC, 1989. Technical report N°34, Methylene chloride (dichloromethane): An overview of experimental work investigating species differences in carcinogenicity and their relevance to man.
- ECETOC, 1990. A Guide to Classification of Preparations Containing Carcinogens, Mutagens and Teratogens. Technical Report N°38.
- ECETOC, 1995. Assessment factors in human health risk assessment. Technical Report No. 68.
- ECETOC, 1996. Practical concepts for dose selection in chronic toxicity and carcinogenicity studies in rodents. Monograph N°25.
- EDO1 Study, 1979. Innovations in cancer risk assessment. Eds. Staffa JA and Mehlman MA Pathotox Publishers Inc.
- Gaylor DW, 1989. Quantitative Risk Analysis for quantal reproductive and developmental effects. *Environ Health Perspect* 74, 243-246.
- Gaylor GW and Chen JJ, 1986. Relative potency of chemical carcinogenesis in rodents. *Risk Anal* 6, 283-290.

- Gold LS, Wright C, Berstein L and De Venciana M, 1987. Reproducibility of results in "near-replicate" carcinogenesis bioassays. *J Nat Cancer Inst* 78, 1149-1158.
- Gold LS, Slone TH and Bernstein L, 1989. Summary of carcinogenic potency and positivity for 492 rodent carcinogens in the carcinogenic potency database. *Environ Health Perspect* 79, 259-272.
- Gold LS, Manley NB, Slone TH, Garfinkel GB and Ames BN, 1993. The fifth plot of the carcinogenic potency database: Results of animal bioassays published in the general literature through 1988 and by the National Toxicology Program through 1988. *Environ. Hlth. Perspec.* 58 9-14.
- Green S, 1991. Receptor mediated non-genotoxic carcinogenesis: new models for risk assessment? *Human & Expt Toxicol.* 10, 390-395.
- Hanes B and Wedel T, 1985. A selected review of risk models: one hit, multi hit, multi-stage, probit, Weibull and pharmacokinetic. *J of Am Coll Toxicol* 4, 271-278.
- Hartley HO and Sielken RL, 1977. Estimation of safe doses in carcinogenic experiments. *Biometrics* 33, 1-30.
- Haseman JK, 1985. Issues in carcinogenicity testing: Dose selection. *Fundam Appl Toxicol* 5, 66-78.
- Hoel DG, 1980. Incorporation of background in dose-response models. *Fed Proc* 39, 73-75.
- Hoel DG, 1982. Statistical measures of risk. *Drugs metab Rev* 13, 829.
- Hoel DG, 1991. A balanced approach to Risk Assessment. *Toxicology and Industrial Health*, 7 - 305-311.
- Hoel DG, Haseman JK, Hogan MD, Huff J and McConnell EE, 1988. The impact of toxicity on carcinogenicity studies: implications for risk assessment *Carcinogenesis* 9, 2045-2052.
- IARC (International Agency for Research on Cancer), 1980. IARC monographs on the Evaluation of Carcinogenic Risks, Suppl 2, Long and Short Term screening Assays for Carcinogens, IARC Lyon, France.
- IARC (International Agency for Research on Cancer), 1987. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans: Suppl 7, IARC, Lyon, France.
- Johannsen FR, 1990. Risk assessment of carcinogenic and noncarcinogenic chemicals. *Critical Reviews in Toxicology*, 20, 341-367.
- Krewski D and Van Ryzin J, 1981. Dose response models for quantal response toxicity data, in *Statistics and Related Topics*. Eds M Csörgő, DA Dawson, JNK Rao and AK Md E, Salen, Pub North-Holland, 201-231.
- Krewski D, Crump KS, Farmer J, Gaylor DW, Howe R, Potier C, Salsburg D, Sielken RL and Van Ryzyn J, 1983. A comparison of statistical methods low dose extrapolation utilising time-to-tumour data. *Fund Appl Toxicol* 3, 140-160.
- Krewski D, Brown C and Murdoch D, 1984. Determining 'safe' levels of exposure: safety factors or mathematical models? *Fundam App Toxicol* 4, S383 - S394.
- Krewski D, Gaylor D and Szyszkowicz M, 1991. A model-free approach to low-dose extrapolation. *Environ Health Perspec* 90, 279-285.
- Kuiper-Goodman T, 1991. Risk Assessment to humans of mycotoxins in animal-derived food products. *Vet Hum Toxicol* 33, 325-333.
- Lu FC and Sielken RL, 1991. Assessment of safety/risk of chemicals; inception and evolution of the ADI and dose-response modelling procedures. *Toxicology Letters* 59, 5-40.
- MAK, 1991. Maximum Concentrations at the Workplace and Biological Tolerance Values for Working Materials Report No 27. Commission for the Investigation Health Hazards of Chemical

compounds in the Work Area. DFG pub VCH, Weinheim.

Mantel N and Bryan WR, 1961. Safety testing of carcinogenic agents. *J Natl Cancer Inst* 27, 455-470.

Moolgavkar SH and Leubeck G, 1990. Two-event model for carcinogenesis: biological, mathematical and statistical considerations. *Risk Anal* 10, 323-341.

Munro IC and Krewski DR, 1981. Risk assessment and regulatory decision making. *Fd Cosmet Toxicol* 19, 549-560.

NTP (National Toxicology Programme), 1984. Report of the NTP *ad hoc* panel on chemical carcinogenesis testing and evaluation. Department of Health and Human Services.

Paustenbach OJ, 1990. Health Risk Assessment and the Practise of Industrial Hygiene. *Am Ind Hyg Assoc J* 51, 339-357.

Paustenbach DJ, Clewell HJ, Gargas ML and Andersen ME, 1987. Development of a physiologically based pharmacokinetic model for multiday inhalation of carbon tetrachloride. In: *Pharmacokinetics in Risk Assessment. Drinking Water and Health*, National Academy Press, Washington DC, 112.

Paynter OE, Burin GJ, Jaeger RB and Gregorio CA, 1988. Goitrogens and thyroid follicular cell neoplasia: Evidence for a threshold procedure. *Regul Toxicol Pharmacol* 8, 102-119.

Peto J, 1985. Some problems in dose-response estimation in cancer epidemiology, In *Methods for Estimating Risk of Chemical Injury: Human and Non-Human Biota and Ecosystems*. Eds VB Honk, GC Butter, DG Hoel and DB Peakal, Pub Scope, 361-380.

Peto R, Pike MC, Day NE, Gray RG, Lee PN, Parish S, Peto J, Richards S and Wahrendorf J, 1980. Guidelines for simple, sensitive significance tests for carcinogenic effects in long-term animal

experiments, in *Long-Term and Short-Term Screening Assays for Carcinogens: A Critical Appraisal*. Lyon: IARC.

Peto R, Gray R, Brantom P and Grasso P, 1991a. Effects on 4080 rats of chronic ingestion of N-nitrosodiethylamine or N-nitrosodimethylamine: a detailed dose response study. *Cancer Res* 51, 6415-6451.

Peto R, Gray R, Brantom P and Grasso P, 1991b. Dose and time relationships and tumour induction in the liver and esophagus of 4080 inbred rats by chronic ingestion of N-nitrosodiethylamine and N-nitrosodimethylamine. *Cancer Res*, 51, 6452-6469.

Planel H, Soleilhavoup JP, Tixador R, Richoilley G, Conter A, Croute F, Caratero C and Gaubin Y, 1987. Influence on cell proliferation of background exposure to very low, chronic gamma radiation. *Health Phys* 59, 11-15.

Portier CJ, Hedges JC and Hoel DG, 1984. Age-specific models and tumour onset for historical control animals in the National Toxicology Program's carcinogenicity experiments. *Cancer Res* 46, 4372-4378.

Rieth JP and Starr TB, 1989a. Chronic bioassays relevance to quantitative risk assessment of carcinogens. *Regul Toxicol Pharmacol* 10, 160-173.

Rieth JP and Starr TB, 1989b. Experimental design constraints on carcinogen potency estimates. *J Toxicol Environ Health* 27, 287-291.

Renwick AG, 1991. Safety factors and establishment of acceptable daily intakes. *Food Add Contam* 8(2), 135-150.

Schuller HM, 1991. The signal transduction model of carcinogenesis. *Biochem Pharm* 42(8), 1511-1523.

Sielken RL, 1987. Cancer dose-response extrapolations. *Environ Sci Technol* 21, 1033-1039.

- Sielken RL, Bretzlaff RS and Stevenson DE, 1995. Challenges to default assumptions stimulate comprehensive realism as a new tier in quantitative risk assessment. *Regul Toxicol Pharmacol* 21, 270-280.
- Swenberg JA, Short B, Borghoff S, Strasser J and Charbonneau M, 1989. The comparative pathobiology of alpha-2 μ -globulin nephropathy. *Toxicol Appl Pharmacol* 97, 35-46.
- UK-DOH (UK Department of Health), 1991. Guidelines for the evaluation of chemicals for carcinogenicity. N°42, HMSO London.
- US-EPA, 1986. Guidelines for Carcinogen Risk Assessment. *Fed Regist*, 33992-34003.
- US-EPA, 1991a. Alpha-2m-Globulin: Association with chemically induced renal toxicity and neoplasia in the male rat. EPA/625/3-91/019F.
- US-EPA, 1991b. The report of the EPA Peer Review workshop on Alpha-2m-globulin Association with Renal Toxicity and Neoplasia in the Male Rat. EPA/625/3-91/02.
- US-EPA, 1991c. Guidelines for developmental toxicity risk assessment. *Fed Reg* 56, 63798-63826.
- US-EPA, 1992. Draft Report : A Cross Species Scaling Factor for Carcinogen Risk Assessment Based on Equivalence of mg/kg 3/4/day. *Fed Reg* 57, 24152-24173.
- Weibull, 1951. Statistical distribution function of wide applicability. *J Appl Mech* 18, 293-302.
- Weisburger EK, 1990. Mechanistic considerations in chemical carcinogenesis. *Regul Toxicol and Pharmacol* 12, 41-52.
- WHO, 1987. Principles for the safety assessment of food additives and contaminants in food. *IPCS Environ Health Criteria* 70, WHO Geneva.
- WHO, 1990. Principles for the toxicological assessment of pesticides in food. *IPCS Environ Health Criteria*, 104. WHO Geneva.
- Wolff S, 1989. Are radiation-induced effects hormetic? *Science* 245, 375.
- Zeise L, Wilson R and Crouch GAC, 1987. Dose response relationship for carcinogens: A review. *Environ Health Persp* 73, 259-308.

MEMBERS OF TASK FORCE

G. Pigott	ZENECA GB - Macclesfield
H. Müsch	KALI-CHEMIE AG D - Hannover
S.A. Hubbard	BORAX EUROPE GB - Wines
B. van Ravenzwaay	BASF D - Ludwigshafen
M. Richold	UNILEVER GB - Sharnbrook
D.D. Lovell	BIBRA GB - Carshalton
D.A. Stringer (until 01.07.95)	ECETOC
W. Rozenboom (from 01.07.95)	B - Brussels

The Task Force wishes to thank Dr. J.P. Rieth (Rhône-Poulenc, USA) for his valuable contribution to this report.

MEMBERS OF THE SCIENTIFIC COMMITTEE

(Peer Review Committee)

W.F. Tordoir (Chairman), Group Adviser, Environmental Health and Human Toxicology	SHELL INTERNATIONAL NL - Den Haag
H. Verschuuren ¹ (Vice-Chairman) Head, Toxicology Department	DOW EUROPE CH - Horgen
O.C. Bockman, Scientific Adviser	NORSK HYDRO N - Porsgrunn
N.G. Carmichael ¹ , Toxicology Director, Worldwide	RHÔNE-POULENC F - Sophia Antipolis
H. De Henau, European Technical Centre, Professional and Regulatory Services	PROCTER AND GAMBLE B - Strombeek-Bever
A. de Morsier, Head, Chemicals Legislation Services	CIBA-GEIGY CH - Basel
C. d'Hondt, Head of Ecology Department	CIBA-GEIGY CH - Basel
P.A. Gilbert, Head, Environmental Division	UNILEVER GB - Port Sunlight
B. Hildebrand, Director, Experimental Toxicology	BASF AG D - Ludwigshafen
J.R. Jackson, Director, Medicine and Health Science	MONSANTO EUROPE B - Louvain-la-Neuve
E. Löser, Head, Institute of Industrial Toxicology	BAYER D - Wuppertal
R. Millischer, Head, Industrial Toxicology Department	ELF ATOCHEM F - Paris
I.F.H. Purchase ¹ , Director, Central Toxicology Laboratory	ZENECA GB - Macclesfield
G. Randall, Director, Brixham Environmental Laboratory	ZENECA GB - Brixham
H.J. Wiegand, Head, Product Safety Department	HÜLS D - Marl

¹ Stewards responsible for primary peer review