

**OECD GUIDANCE DOCUMENT ON THE DESIGN AND CONDUCT OF CHRONIC TOXICITY  
AND CARCINOGENICITY STUDIES, SUPPORTING TG 451, 452 AND 453**

**TABLE OF CONTENTS**

**NOTE: Only drafts of Chapter 1, General Introduction, Chapter 2, Guidance on Developing a Mode of Action and section 3.1 of Chapter 3, Dose Selection, of this draft Guidance Document are currently available. This draft was developed in parallel with the revision of Test Guidelines 451, 452 and 453 on carcinogenicity, chronic toxicity and combined carcinogenicity/chronic toxicity studies. Additional chapters will be developed and the Guidance Document will be updated.**

PREFACE, ABOUT THE OECD

	<b>PAGE</b>
1. GENERAL INTRODUCTION	
1.1 Guiding principles and considerations	3
1.2 Scope of application of the guidance	4
1.3 Objectives of a chronic toxicity study	5
1.4 Objectives of a carcinogenicity study	5
1.5 Objectives of a combined chronic toxicity/carcinogenicity study	5
1.6 Principles of intelligent testing	6
1.7 Animal welfare considerations	8
REFERENCES	9
2. GUIDANCE ON DEVELOPING A MODE OF CARCINOGENIC ACTION	
2.1 Introduction	12
2.2 “Mode of Action” Definition	13
2.3 Mode of Action Framework: Animal Tumours	13
2.4 Components of a Mode of Action Analysis	13
2.5 Postulated Mode of Action: Key Events	14
2.6 Experimental Support for the Postulated Mode of Action	14
2.7 Alternative Modes of Action	15
2.8 Uncertainties, Inconsistencies and Data Gaps	15
2.9 Conclusion of Postulated Mode of Action Analysis	15
2.10 Relevance of Rodent Mode of Action for Humans	15
2.11 Human Relevance Framework	16
2.12 Hazard Characterisation	16
2.13 More information	17
References	17
Appendix 1: Examples of Animal Mode(s) of Action Framework	19
Appendix II: Example of Human Relevance Framework	20
3. STUDY DESIGN	21
3.1 Dose Selection	21
3.1.1 Introduction	21
3.1.2 Objectives of a Chronic Bioassay	23
3.1.2.1 Key information for the selection of doses in chronic toxicity and carcinogenicity studies	23
3.1.2.2 Selection of the highest dose level	25
3.1.2.3 Dose level spacing	28
3.1.2.4 Relevance of effects at high dose levels	29

3.1.3	Objectives of a chronic bioassay and their influence on dose selection	29
3.1.3.1	Screening Chemicals to Identify Carcinogens (or other toxic effects)	30
3.1.3.2	Characterizing the Dose–Response Curve in the Observable Range	31
3.1.3.3	Characterizing the Dose–Response Curve to Facilitate Low-Dose (Linear) Extrapolation	31
3.1.3.4	Defining a Threshold or BMD Departure Point	32
3.1.3.5	Providing Data on Health Effects at Human Exposure Levels	33
3.1.3.6	Providing Data to Test Hypotheses Regarding Mode of Action of a Carcinogen	34
3.1.3.7	Integration of the objectives of a chronic bioassay	34
	References	36
	Appendix 1: ILSI Principles for Dose Selection in Chronic Rodent Bioassay	38
3.2	Route of Exposure and Dose Administration Considerations	
3.2.1	Oral	
3.2.2	Dermal	
3.2.3	Inhalation	
3.2.4	Other Routes	
3.3	Choice of Species and Strain, Numbers & Sex of Animals, Study Duration	
3.3.1	Testing in rodents	
3.3.1.1	Rodent species and strain specificity	
3.3.2	Study duration	
3.3.3	Testing in non-rodents	
3.3.4	Alternative models for carcinogenicity testing, including testing in transgenic animals	
3.4	Toxicokinetics	
3.5	Housing, Feeding, Handling of Animals and Experimental Procedures	
3.6	Investigations (Including Histopathological Guidance)	
3.7	Statistical and Dose Response Analysis, Including Benchmark Dose and Linear Extrapolation, NOAELS and NOELS, LOAELS and LOELS	
4.	DOCUMENTATION AND DATA ACCEPTABILITY, USE OF HISTORICAL CONTROL DATA	
5.	REPORTING	
6.	DEFINITIONS/GLOSSARY	

## 1. GENERAL INTRODUCTION

### 1.1 Guiding principles and considerations

1. Chronic toxicity and carcinogenicity studies aim to determine toxic effects and potential health hazards following prolonged, repeated exposure. This type of study is usually required if humans are likely to be exposed to a substance over a significant portion of their life span. In the 1960s, long-term (chronic) animal bioassays began to be routinely used for hazard identification, to assess the qualitative potential of a chemical to cause chronic toxicity and cancer.

2. The objectives of the chronic bioassay have however expanded beyond hazard identification and are now focused primarily on assessment of risk for humans. In addition there has been increasing pressure for the chronic bioassay design to consider financial constraints and societal desires to minimize the number of animals needed for scientific interpretation of results.. There is a growing desire for chronic studies to provide data that cover a number of objectives including characterization of the nature of specific toxic responses, description of dose–response relationship, establishment of inflection points, and provision of insight into the roles of toxicokinetics and mechanisms of toxic action. In practice, it is likely that the bioassay design will be a compromise among a set of different purposes; to the extent that the ability to address one question is enhanced, the ability to address others may be diminished. For example, it may be necessary to achieve a balance between the power to detect toxicity and the ability to estimate the dose–response relationship of any observed effects.

3. The use of formal risk assessment procedures by government regulatory bodies began to emerge in the late 1970s and early 1980s bringing with it a strong interest in using data for quantitative as well as qualitative purposes. The need to gather data that allowed an understanding of the shape and slope of the dose-response curve focused attention on the number of doses in a bioassay and their spacing. Advances in knowledge of how chemicals perturbed or otherwise modulated biological processes in the development of tumors or other forms of toxicity provided bases for further improving the risk assessment process. Through meetings held primarily under the auspices of the International Programme on Chemical Safety (IPCS), a Mode of Action (MoA) framework (3) was developed and refined (Sonnich-Mullin et al., 2002; Cohen et al. 2003; Meek et al, 2003; Holsapple et al, 2006; Boobis et al., 2006, EPA, 2005), as outlined further in Chapter 2 of this guidance. The key purpose of this work was to introduce greater transparency into the process of assessing human relevance, and the goal was to use a broad array of relevant data to determine the predictive value of a bioassay tumor response to risk in humans.

4. The broadened range and complexity of scientific data used to evaluate chemical toxicity and carcinogenicity potential for humans highlighted the need to revise and update OECD Chronic Toxicity and Carcinogenicity Test Guidelines 451, 452 and 453, originally adopted in 1981. These Guidelines have therefore recently been revised in the light of scientific progress and the updating of related OECD Guidelines such as TG 408 (90-day oral toxicity study in rodents) and TG 407 (28-day oral toxicity study in rodents).

5. During the revision of the Test Guidelines, an emphasis was placed on providing guidance on factors that influence the selection of test doses, particularly for carcinogenicity studies. It was recognized that while general principles of dose selection should be contained in the Test Guidelines themselves, there was a need for additional guidance on these principles. These considerations were influenced by the publication of two reports by the International Life Sciences Institute (ILSI), “*Principles for the Selection of Doses in Chronic Rodent Bioassays*” (ILSI, 2007), and “*Issues in the Design and Interpretation of Chronic Toxicity and Carcinogenicity Studies in Rodents: Approaches to Dose Selection*” (Rhomberg et al., 2007) These

reports provided theoretical and practical guidance on factors that influence dose selection in chronic bioassays.

6. A summary of the principles contained in these two publications, to underpin the texts on dose selection contained in the Test Guidelines is provided in this guidance (in chapter 3.1). During the development of this material, suggestions were made for additional guidance on specific aspects of study design in relation to core objectives of these studies, and how they might impact on other aspects of the study (e.g. design for optimising carcinogenicity data versus chronic toxicity, design of studies for risk rather than hazard assessment) and on statistical power. It was generally agreed that the scope of the guidance should be wider than principles of dose selection, and should cover a number of key issues related to carcinogenicity and chronic toxicity testing, including the need for pharmacokinetic and mechanistic data.

7. This guidance therefore provides additional information on the conduct of studies performed using TG 451, 452 and TG 453. Its objective is to assist users of the Test Guidelines to select the most appropriate methodology to assess the chronic toxicity and carcinogenicity of a test chemical so that particular data requirements can be met while reducing animal usage and suffering. It should be noted, however, that the basic principles for the conduct of chronic toxicity and carcinogenicity studies will differ, given that the endpoints are different. While the guidance provided in this document can be taken as generally applicable to the conduct of either a chronic toxicity or a carcinogenicity bioassay, or a combined chronic and carcinogenicity study, users of the guidance should be mindful of the primary objectives of the study (8) in establishing the optimum study design.

8. The guidance aims to foster a common approach among those carrying out chronic toxicity and carcinogenicity studies, and thereby contribute to the harmonisation activities undertaken by the OECD and other agencies, such as the WHO<sup>1</sup>. It should be consulted alongside other national guidance and requirements documents. It provides broad guidance on approaches to the execution of chronic toxicity and carcinogenicity studies, and on some of the problems and pitfalls that may arise during an assessment of possible compound-related changes in parameters measured in toxicity studies. The text reflects scientific understanding and standards as at the date of issue. In time, the scientific community will gain a better understanding of the mechanisms of toxicity, and this may lead to changes in both methodology and interpretation of results; which should reflect scientific consensus at the time data are reviewed.

9. Two other OECD documents provide guidance on the analysis and evaluation of the results of repeat-dose toxicity studies and chronic toxicity and carcinogenicity studies, No. 32, Guidance Notes for Analysis and Evaluation of Repeat-Dose Toxicity Studies (OECD, 2002a) and No. 35, Guidance Notes for Analysis and Evaluation of Chronic Toxicity and Carcinogenicity Studies (OECD, 2002b).

## 1.2 Scope of application of the guidance

10. The Test Guidelines 451, 452 and 453 are designed to be used in the testing of a wide range of chemicals, including pesticides, industrial chemicals and pharmaceuticals. This guidance is similarly intended to provide information generally applicable to the testing of a wide range of chemicals, whatever their field of application. However, as noted in the Test Guideline 451, some testing requirements may differ for pharmaceuticals. The International Conference on Harmonisation of Testing for Pharmaceuticals (ICH) has produced a series of safety guidelines for the testing of pharmaceuticals, including guidelines on

---

<sup>1</sup> An example, already referred to, is the IPCS project *Harmonisation of Approaches to the Assessment of Risk from Exposure to Chemicals*, which has developed a *Conceptual Framework for Cancer Risk Assessment*. The framework is an analytical tool for judging whether the available data support a postulated mode of carcinogenic action.

toxicokinetics (S3A), genotoxicity testing (S2), duration of chronic toxicity testing in animals (rodent and non-rodent toxicity testing) (S4), testing for carcinogenicity of pharmaceuticals (S1B) and dose selection for carcinogenicity studies of pharmaceuticals (S1C). These Guidelines should always be consulted for specific guidance when testing pharmaceuticals using the approaches outlined in TG 451, 452 and 453, although this guidance document provides, in various sections, examples of where the testing requirements may be different for pharmaceuticals

### **1.3 Objectives of a chronic toxicity study**

11. The objective of a chronic toxicity study is to characterise the profile of a substance in a mammalian species following prolonged and repeated exposure. The chronic oral toxicity study provides information on the possible health hazards likely to arise from repeated exposure over a prolonged period of time covering post-weaning maturation and growth well into adulthood. A key objective of the study is to provide an estimate of the no-observed-adverse-effect level (NOAEL) of exposure which can be used for establishing safety criteria for human exposure. The study will provide information on the major toxic effects, indicate target organs and the possibility of accumulation. The need for careful clinical observations of the animals, so as to obtain as much information as possible, is also stressed. Previous repeated dose 28-day and/or 90-day toxicity tests on a chemical may have indicated the potential to cause neurotoxic/neurobehavioural effects, warranting further in-depth investigation as part of a chronic oral toxicity study.

12. The duration of chronic toxicity studies for effects other than neoplasia is still widely debated. Under the conditions of this test, effects such as carcinogenesis and those which are not specifically life shortening, which require a long latent period, or are cumulative may not become manifest. Except for those effects, the application of these Guidelines should generate data on which to identify the majority of chronic effects and to determine dose-response relationships. Ideally, the design and conduct should allow for the detection of general toxicity including neurological, physiological, biochemical and haematological effects and exposure-related morphological (pathology) effects.

### **1.4 Objectives of a carcinogenicity study**

13. The objective of a long-term carcinogenicity study is to observe test animals for a major portion of their life span for the development of neoplastic lesions during or after exposure to various doses of a test substance by an appropriate route of administration. The carcinogenicity study provides information on the possible health hazards likely to arise from repeated exposure for a period lasting up to the entire lifespan of the species used. The study will provide information on the toxic effects of the substance including potential carcinogenicity, and may indicate target organs and the possibility of accumulation. It can provide an estimate of the no-observed-adverse effect level for toxic effects and, in the case of non-genotoxic carcinogens, for tumour responses, which can be used for establishing safety criteria for human exposure. The need for careful clinical observations of the animals, so as to obtain as much information as possible, is also stressed. Such an assay requires careful planning and documentation of the experimental design, a high standard of pathology, and unbiased statistical analysis. These requirements are well known and have not undergone any significant changes in recent years.

### **1.5 Objectives of a combined chronic toxicity/carcinogenicity study**

14. The objective of a combined chronic toxicity/carcinogenicity study is to determine the effects of a substance in a mammalian species following prolonged and repeated exposure. The combined chronic toxicity/carcinogenicity study provides information on the possible health hazards likely to arise from repeated exposure over the majority of the entire lifespan (in rodents). The study will provide information on the toxic effects of the substance including potential oncogenicity, indicate target organs and the

possibility of accumulation. The need for careful clinical observations of the animals, so as to obtain as much information as possible, is also stressed. In conducting such a study, the guiding principles and considerations outlined in the OECD Guidance Document on the recognition, assessment, and use of clinical signs as humane endpoints for experimental animals used in safety evaluation (14), in particular paragraph 62 thereof, should always be followed. The application of the Test Guideline 453 should generate data on which to identify the majority of chronic and carcinogenic effects and to determine dose-response relationships. Ideally, the design and conduct should allow for the detection of neoplastic effects and a determination of carcinogenic potential as well as general toxicity, including neurological, physiological, biochemical, and haematological effects and exposure-related morphological (pathology) effects.

15. The design of the updated test guideline TG 453 recommends, for the chronic phase of the study, at least three dose groups and a control group, each group containing at least 10 males and 10 females per group, as in the chronic toxicity guideline TG 452. Similarly, the design of the carcinogenicity phase in the two guidelines is identical: The study will thus provide similar information on chronic toxicity and carcinogenicity as TG 452 and TG 451. It will allow derivation of a NOAEL where relevant and a BMD (if needed), and will offer greater efficiency in terms of time and cost compared to conducting two separate studies, without compromising the quality of the data in either the chronic phase or the carcinogenicity phase. The two phases (chronic toxicity and carcinogenicity) will reinforce each other in terms of the two pools of animals on study, since they will have been drawn from the same stock and have similar characteristics at the start of the study. Measurements carried out on the animals in one phase will be relevant for the animals in the other phase, e.g. clinical signs, body weights, haematology and biochemistry (if carried out), pathology. The terminal kill of the chronic phase can act as an interim kill for the carcinogenicity phase, offering some saving in animals. It should be noted, however, that the savings in animals is not substantial, and that the study may be more complex to execute.

## **1.6 Principles of intelligent testing**

16. This guidance document does not recommend any particular testing strategy or approach, but suggests consideration of such approaches as part of an ongoing strategy to assess the toxic potential of a substance in an intelligent and iterative manner. As new methods or approaches become scientifically appropriate for use in chronic toxicity or carcinogenicity assessment, the study director is encouraged to implement them where possible.

17. A reasoned scientific approach to the assessment of substances for chronic toxicity or carcinogenicity must first include an assessment of all available information that has the potential to influence the study design. This can include the identity, molecular structure, class, and physico-chemical properties of the test substance; any information regarding mode of action; results of relevant *in vitro* or *in vivo* toxicity tests such as genotoxicity, subchronic toxicity and toxicokinetics; anticipated use(s) and potential for human exposure; available (Q)SAR data; and relevant toxicological data on structurally-related substances. This analysis can focus the study parameters, but may also lead to the conclusion that a study can be refined in some way, or not conducted at all (Carmichael et al 2006, Doe et al 2006, Barton et al 2006, Cooper et al 2006).

18. Integrating a wide range of information to determine the potential toxicity of a substance is becoming more common as the gap between assessments that need to be conducted and the resources with which to conduct such assessments widens. Efforts are underway in many OECD countries to determine ways in which assessments of substances can be satisfactorily completed, and protection of public health and the environment achieved, while minimising costs in terms of time, money and animal use. However, the acceptability and use of testing strategies and weight-of-evidence approaches differ among OECD

countries and regulatory sectors; thus, application of these approaches should always occur in consultation with appropriate regulatory authorities.

19. In general, testing strategies are conducted in a phased or step-wise manner, beginning with database mining, (Q)SAR modelling, and high- to medium-throughput *in vitro* tests, followed by short-term *in vivo* studies, and finally, if warranted, long-term tests. At each step, the available evidence is evaluated to determine what additional testing, if any, is needed to make a regulatory decision, in order to avoid studies that would provide no additional information for that particular purpose (van Leeuwen et al 2007). Testing strategies are a natural extension of weight-of-evidence principles that rely on informed scientific judgment to determine the potential chronic toxicity or carcinogenicity of a test substance (OECD 2002a).

20. As stated above, other shorter-term *in vitro* or *in vivo* tests may provide information regarding potency, mode of action, metabolism, and/or target organ that can help refine the chronic toxicity study protocol parameters or priorities for observation. Tiered approaches using a combination of *in silico*, *in vitro*, and *in vivo* tests have been proposed but are not yet widely implemented (Becker et al 2007; Worth and Balls 2002).

21. A phased or tiered approach to the assessment of the carcinogenic potential of a substance should also be considered (Ashby 1996). A number of shorter-term tests can be conducted which will provide information of use in determining whether and how a substance may be carcinogenic, including genetic toxicity assays, cell transformation or other cell-based assays, short-term cancer initiation-promotion tests which may or may not include toxicogenomic analyses (Ellinger-Ziegelbauer et al 2005; 2008), and *in vivo* repeated dose 28- or 90-day toxicity tests (for a review see Maurici et al 2005). (Q)SAR prediction models have been used in a regulatory context to predict the carcinogenic potential of chemicals for several decades; a recent survey found 16 non-commercial models available for predicting rodent carcinogenicity (Begnini et al 2007).

22. In most circumstances, a carcinogenicity test would not be carried out on a known *in vivo* genotoxicant, but this is not always the case. Potential exceptions include compounds of high value or substances with unavoidable exposures. Also, there may be mechanistic data to suggest that the genotoxicity will not be reflected in a carcinogenic hazard.

23. The US National Toxicology Program, along with agencies in other OECD countries, has had a longstanding interest in the use of transgenic or knockout mouse models for the assessment of carcinogenicity (Bucher and Portier 2004), as they consider that these models offer potential refinements, in terms of study duration and animal numbers, over the traditional long-term bioassay. Currently, some regulatory authorities in the pharmaceutical sector may accept studies with these models in combination with a full long-term rat bioassay in lieu of a second full bioassay in mice (See ICH 1997). However, the predictive ability of the models, and any refinements or animal reductions, have fallen short of expectations (Goodman 2001; van Zeller and Combes 1999). A draft detailed review paper (DRD) on transgenic rodent mutation assays prepared by Canada is currently under discussion at OECD.

24. In the case of carcinogenicity or chronic toxicity testing by the dermal route of exposure, consideration should be given as to whether the substance is absorbed appreciably through the skin. In cases where dermal bioavailability is low, testing via the dermal route may not be warranted.

25. A number of different strategies for assessing carcinogenicity have been proposed that take advantage of some of the tests mentioned here (Worth and Balls 2002; Langley 2001; Knight et al 2006; Combes et al 2008). All feature a stepwise process or decision tree that prescribes information analysis and stopping points where classification and labelling and/or risk assessment could be possible. However, specific approaches have not yet been optimised or validated.

26. Consideration of particular tests or approaches should always be made within the context of whether the results will contribute mechanistic information that will be useful in the weight-of-evidence assessment of carcinogenic potential (OECD 2002b).

### **1.7 Animal welfare considerations**

27. The principles of the “3Rs” (Replacement, Reduction, and Refinement), first articulated by Russell and Birch in 1959 (Russell and Birch 1959), should be considered as integral to the assessment of carcinogenicity or chronic toxicity in mammals, in order to ensure sound science, maximize animal welfare, and minimize animal use. Animals in a condition of stress or distress have a documented effect on the outcome of the study (Olsson and Dahlborn 2001; Reinhardt and Reinhardt 2002; and references therein). For these reasons the following principles should be implemented as much as practicably possible.

28. First and foremost, as discussed above, consideration of existing information from any source that could provide a refinement in the testing protocol or procedure is recommended. Existing information could be used to inform dose spacing or selection, exposure route, observation priorities, potential modes of action or target organs of the test substance, and/or study design. Use of this information to focus the study before it begins ensures that the study will meet the expectations of the study director or regulatory authorities, decreasing the likelihood of repeat studies.

29. The use of the combined chronic toxicity/carcinogenicity study (TG453) is also recommended, which can in most cases accomplish the objectives of both studies, and offers significant savings in the numbers of animals used.

30. Any studies involving animals should abide by the principles of humane euthanasia as detailed in the OECD Guidance Document 19 on the recognition, assessment, and use of clinical signs as humane endpoints for experimental animals used in safety evaluation, and in particular paragraph 62 thereof (OECD 2000). This paragraph states that “In studies involving repeated dosing, when an animal shows clinical signs that are progressive, leading to further deterioration in condition, an informed decision as to whether or not to humanely kill the animal should be made. The decision should include consideration as to the value of the information to be gained from the continued maintenance of that animal on study relative to its overall condition. If a decision is made to leave the animal on test, the frequency of observations should be increased, as needed. It may also be possible, without adversely affecting the purpose of the test, to temporarily stop dosing if it will relieve the pain or distress, or reduce the test dose.” Close and frequent observations are recommended in order to determine the status of the animals, and any animals exhibiting clear signs of severe pain or distress should be humanely killed.

31. Housing in small groups of the same sex is recommended; individual housing must be scientifically justified. It is well documented that group housing provides numerous animal welfare benefits, especially in chronic or lifetime-length studies. Further detailed information on housing, feeding, and substance administration is provided in Chapter 3.5.

32. The selection of dosing route, level, and spacing in accordance with the objectives of the study is discussed in detail in Chapter 3. The Maximum Tolerated Dose concept, while still practiced, is not favoured for both practical and animal welfare reasons, and care should be taken to avoid the selection of overly high doses or dosing above the prescribed limit dose (normally 1000 mg/kg), while also balancing the need to ensure the acceptability of a negative study outcome. Selection of such high doses must be justified.

33. As discussed further in section 3.2 of chapter 3, while the route of administration will depend on the physical and chemical characteristics of the test substance and expected route of human exposure, mixing the test substance into the diet or water is normally recommended. Administration of the test substance by oral gavage in carcinogenicity and chronic toxicity testing is normally not recommended for the reasons outlined in section 3.2. The testing of potentially irritating or corrosive substances should be avoided, as administering such substances could result in severe pain and tissue damage at point-of-entry, which would compromise both animal welfare and the integrity of the study.

34. Testing the chronic toxicity or carcinogenicity of inhaled substances can be achieved using either of two exposure conditions: whole-body or nose-only/snout-only. Because of the long-term nature of chronic toxicity and carcinogenicity tests, the preferred exposure method is whole-body. Deviations from this recommendation, especially for smaller animals, should be scientifically justified.

## REFERENCES

Ashby J (1996) Alternatives to the 2-species bioassay for the identification of potential human carcinogens. *Hum Exp Toxicol* 15: 183-202.

Barton HA, Pastoor TP, Baetcke T, et al (2006) The Acquisition and Application of Absorption, Distribution, Metabolism, and Excretion (ADME) Data in Agricultural Chemical Safety Assessments. *Crit Rev Toxicol* 36: 9-35.

Becker, RA, Plunkett, LM, Borzelleca, JF, & Kaplan, AM. (2007) Tiered toxicity testing: Evaluation of toxicity-based decision triggers for human health hazard characterization. *Food and Chemical Toxicology* 45: 2454–2469.

Benigni R, Netzeva TI, Benfenati E, Bossa C, Franke R, Helma C, Hulzebos E, Merchant C, Richard A, Woo YT & Yang C (2007) The Expanding Role of Predictive Toxicology: An Update on the (Q)SAR Models for Mutagens and Carcinogens. *J Env Sci Hlth Part C*, 25: 53-97.

Boobis, A.R., Cohen, S.M., Dellarco, V., McGregor, D., Meek, M.E., Vickers, C., Willcocks, D. & Farland, W. IPCS Framework for analyzing the Relevance of a Cancer Mode of Action for Humans. *Crit. Rev. in Toxicol*, (2006) 36:793-801.

Bucher J & Portier C (2004) Human Carcinogenic Risk Evaluation, Part V: The National Toxicology Program Vision for Assessing the Human Carcinogenic Hazard of Chemicals. *Tox Sci* 82: 363–366.

Cohen, S.M., Meek, M.E., Klaunig, J.E., Patton, D.E., and Fenner-Crisp, P.A. (2003) The human relevance of information on carcinogenic Modes of Action: An Overview. *Crit. Rev. Toxicol.* 33:581-589

Cooper RL, Lamb JS, Barlow SM, et al (2006) A Tiered Approach to Life Stages Testing for Agricultural Chemical Safety Assessment. *Crit Rev Toxicol* 36: 69-98.

Carmichael NG, Barton HA, Boobis AR, et al (2006) Agricultural Chemical Safety Assessment: A Multisector Approach to the Modernization of Human Safety Requirements. *Crit Rev Toxicol* 36: 1-7.

Combes R, Grindon C, Cronin MTD, Roberts DW & Garrod JF (2008) Integrated Decision-tree Testing Strategies for Mutagenicity and Carcinogenicity with Respect to the Requirements of the EU REACH Legislation. *Alt Lab Anim* 36: 43-63.

Doe JE, Boobis AR, Blacker A, et al (2006) A Tiered Approach to Systemic Toxicity Testing for Agricultural Chemical Safety Assessment. *Crit Rev Toxicol* 36: 37-68.

Ellinger-Ziegelbauer H, Gmuender H, Bandenburg A & Ahr HJ (2008) Prediction of a carcinogenic potential of rat hepatocarcinogens using toxicogenomics analysis of short-term in vivo studies. *Mutat Res* 637: 23-39.

Ellinger-Ziegelbauer H, Stuart B, Wahle B, Bomann W & Ahr HJ (2005) Comparison of the expression profiles induced by genotoxic and non-genotoxic carcinogens in rat liver. *Mutat Res / Fundamental and Molecular Mechanisms of Mutagenesis* 575: 61-84.

EPA (2005). Guidelines for Carcinogen Risk Assessment Risk Assessment Forum U.S. Environmental Protection Agency Washington, DC

Goodman JI (2001) A perspective on current and future uses of alternative models for carcinogenicity testing. *Toxicol Pathol* 29 (Suppl): 173-176.

Holsapple, M.P., Pitot, H.C., Cohen, S.N., Boobis, A.R., Klaunig, J.E., Pastoor, T., Dellarco, V.L. & Dragan, Y.P. (2006) Mode of Action in Relevance of Rodent Liver Tumors to Human Cancer Risk. *Toxicol. Sci.* 89:51-56.

ICH (International Conference on Harmonisation). 1997. ICH Harmonised Tripartite Guideline - Testing for Carcinogenicity of pharmaceuticals - S1B.  
Available at: <http://www.ich.org/LOB/media/MEDIA490.pdf>.

ILSI (International Life Sciences Institute) (1997). Principles for the Selection of Doses in Chronic Rodent Bioassays. Foran JA (Ed.). ILSI Press, Washington, DC.  
Knight A, Bailey J & Balcombe J (2006) Animal Carcinogenicity Studies: 3. Alternatives to the Bioassay. *Alt Lab Anim* 34: 30-48.

Langley G. (2001) The way forward - Action to end animal toxicity testing, 36pp. London, UK: BUAV.

Maurici D, Aardema M, Corvi R, et al. (2005) Carcinogenicity. *Alt Lab Anim* 33 (Suppl 1): 177-182.

Meek, E.M., Bucher, J.R., Cohen, S.M., Dellarco, V., Hill, R.N., Lehman-McKemmon, L.D., Longfellow, D.G., Pastoor, T., Seed, J. & Patton, D.E. A Framework for Human Relevance analysis of Information on Carcinogenic Modes of Action. *Crit. Rev. Toxicol.* (2003). 33:591-653.

OECD (2000) OECD Series on Testing and Assessment No. 19. Guidance Document on the recognition, assessment, and use of clinical signs as humane endpoints for experimental animals used in safety evaluation. Organisation for Economic Co-operation and Development, Paris. Available at: [http://www.oelis.oecd.org/olis/2000doc.nsf/LinkTo/NT00002E46/\\$FILE/00087372.PDF](http://www.oelis.oecd.org/olis/2000doc.nsf/LinkTo/NT00002E46/$FILE/00087372.PDF).

OECD (2002a) OECD Series on Testing and Assessment No. 32: Guidance Notes for Analysis and Evaluation of Repeat-Dose Toxicity Studies. Organisation for Economic Co-operation and Development, Paris. Available at: [http://www.oelis.oecd.org/olis/2000doc.nsf/LinkTo/NT00002CCE/\\$FILE/JT00129312.PDF](http://www.oelis.oecd.org/olis/2000doc.nsf/LinkTo/NT00002CCE/$FILE/JT00129312.PDF).

OECD (2002b) OECD Series on Testing and Assessment No. 35: Guidance Notes for Analysis and Evaluation of Chronic Toxicity and Carcinogenicity Studies. Organisation for Economic Co-operation and Development, Paris.

Available at:

[http://www.oalis.oecd.org/olis/2002doc.nsf/LinkTo/NT00002BE2/\\$FILE/JT00130828.PDF](http://www.oalis.oecd.org/olis/2002doc.nsf/LinkTo/NT00002BE2/$FILE/JT00130828.PDF). Accessed 3 March 2009.

Olsson A & Dahlborn K. (2001) Improving housing conditions for laboratory mice: a review of 'environmental enrichment'. *Lab Animals* 36: 243-270.

Reinhardt V. & Reinhardt A. (2002) Comfortable quarters for laboratory animals. Animal Welfare Institute, Washington, D.C.

Rhomberg, L., Baetcke, K., Blancato, J., Bus, J., Cohen, S., Conolly, R., Dixit, R., Doe, J., Ekelman, K., Fenner-Crisp, P., Harvey, P., Hattis, D., Jacobs, A., Jacobson-Kram, D., Lewandowski, T., Liteplo, R., Pelkonen, O., Rice, J., Somers, D, Turturro, A., West, W. Olin, S. (2007). Issues in the Design and Interpretation of Chronic Toxicity and Carcinogenicity Studies in Rodents: Approaches to Dose Selection. *Crit. Rev. Toxicol.*37:729-837.

Russell WMS & Burch RL (1959) *The Principles of Humane Experimental Technique*.

Available at: [http://altweb.jhsph.edu/publications/humane\\_exp/het-toc.htm](http://altweb.jhsph.edu/publications/humane_exp/het-toc.htm).

Sonnich-Mullin, C., Fielder, R., Wiltse, J., Baetcke, K., Dempsey, J., Fenner-Crisp, P., Grant, D., Hartley, M., Knaap, A., Kroese, D., Mangelsdorf, I., Meek, E., Rice, J., & Younes, M. IPCS Conceptual Framework for Evaluating a Mode of Action for Chemical Carcinogenesis. *Reg. Toxicol. & Pharm.* (2001) 34:146-152.

Van Leeuwen CJ, Patlewicz GY, and Worth AP. (2007) Intelligent Testing Strategies. In *Risk Assessment of Chemicals, An Introduction 2<sup>nd</sup> Edition*. C.J. van Leeuwen and T.G.Vermeire, eds. Springer Netherlands, Rotterdam, pp 467-509.

Van Zeller AM & Combes RD (1999) Transgenic Mouse Bioassays for Carcinogenicity Testing: A Step in the Right Direction? *Alt Lab Anim* 27: 39–846.

Worth A & Balls M (2002). *Alternative (Non-animal) Methods for Chemicals Testing: Current Status and Future Prospects*. A Report prepared by ECVAM and the ECVAM Working Group on Chemicals. *Alt Lab Anim* 30 (Suppl 1): 1-125.

## 2. GUIDANCE ON DEVELOPING A MODE OF CARCINOGENIC ACTION

(Mary Ko Manibusan, USEPA/Office of Pesticide Programs; Douglas C. Wolf, D.V.M., Ph.D., Fellow IATP, ATS, USEPA/Office of Research and Development and Vicki Dellarco, Ph.D. USEPA/Office of Pesticide Programs)

### 2.1 Introduction

35. Chronic exposure animal studies have been used for more than a half century to determine whether pesticides, pharmaceuticals, industrial chemicals, and other products have the potential to cause cancer or other health problems in humans. As such, these chronic bioassays have become the standard for detecting carcinogenic potential of products where human use or exposure is anticipated. In most cases extrapolation of dose and species is necessary from animal data to predict and estimate human cancer risk. Few agents have sufficient human data upon which to base cancer assessments. Inherent in these animal based assessments are the assumptions that observation of tumors in animals is relevant for human cancer risk and that responses observed at high doses in animals are meaningful to doses of regulatory relevance for humans (IPCS, 2005). Such extrapolations have been surrounded by intense discussion and debate. Data from molecular and cellular studies have resulted in a fuller biological understanding of how chemicals induce neoplasia in animal studies. Such mechanistic work has also raised concern regarding the appropriateness of extrapolating positive tumor responses in rodents to human cancer risk (Holsapple et al., 2005).

36. Significant progress has been made in the last twenty years in our understanding of the biological process(es) whereby carcinogenesis occurs in both animals and in humans (Hanahan and Weinberg, 2000). Risk assessments have improved through the organized scientific use of data on the toxicokinetics and toxicodynamics of agents that characterize the biological basis underlying the use of assumptions. More recently, the U.S. EPA's issuance of its revised Risk Assessment Guidelines for Carcinogen (U.S. EPA, 2005) recognized this advancement in scientific thinking by highlighting the use of mode of action as a framework to test hypothesized toxicity pathways that could lead to a carcinogenic response. Adoption of the mode of action framework is widespread and it is commonly used by many regulatory agencies and international organizations. In the United Kingdom, the mode of action framework is routinely utilized for the assessment of both pesticides and industrial chemicals. The United Kingdom Committee in its latest Guidelines (COC, 2004) on Carcinogenicity has noted the value of the mode of action framework with regard to both harmonization between agencies and internal consistency. This framework is also used by agencies in Australia and Canada. In Canada the mode of action framework is used for the evaluation of Existing Chemicals under the Canadian Environmental Protection Act. The European Union has incorporated the framework into the technical guidance documents that are being updated on evaluating new and existing industrial chemicals and biocides, for carcinogenicity. With regard to the international organizations, the framework is being used by the WHO/FAO Joint Meeting on Pesticide Residues in its evaluation of pyrethrin extract and its incorporation into the resulting monograph. The step to progress this forward to the human relevance concept has been taken by IPCS in cooperation with international partners. (IPCS Workshop, 2005). To refine and improve the process of carcinogenic hazard identification, and to avoid misidentification of harmless substances as possible human carcinogens, it has become imperative that mode of action analysis be the standard analytic approach for regulatory purposes and that data to support such analysis be collected in a thorough and scientifically rigorous manner that informs the mode of action. (Rice, 2004)

*Note: This guidance document is based largely from the International Programme on Chemical Safety (IPCS) Workshop Report (IPCS, 2005 and Boobis et al., 2006) and the U.S. EPA Guidelines for Carcinogen Risk Assessment, (USEPA, 2005).*

## **2.2 “Mode of Action” Definition**

37. The term “mode of action” is defined as a sequence of key events and processes, starting with interaction of an agent with a cell, proceeding through operational and anatomical changes, and resulting in cancer formation. A “key event” is an empirically observable precursor step that is itself a necessary element of the mode of action or is a biologically based marker for such an element. Mode of action is contrasted with “mechanism of action” which implies a more detailed understanding and description of events, often at the molecular level, than is implied by mode of action. The toxicokinetic processes that lead to formation or distribution of the active agent to the target tissue are considered in estimating dose but are not part of the mode of action as the term is used here. There are many examples of possible modes of carcinogenic action, such as mutagenicity, mitogenesis, inhibition of cell death, cytotoxicity with reparative cell proliferation, and immune suppression. (Refer to Appendix I. for examples of mode(s) of action)

38. Elucidation of a mode of action for a particular carcinogenic response in animals or humans is a data rich determination. Significant information should be developed to ensure that a scientifically justifiable mode of action underlies the process leading to cancer at a given site. In the absence of sufficient data for a scientifically defensible mode of action, EPA generally takes a public health protective approach regarding interpretation of data including animal tumor findings. In the case of insufficient data, the animal data are considered relevant to humans and cancer risks are assumed to conform to low dose linearity.

## **2.3 Mode of Action Framework: Animal Tumors**

39. The framework is intended to be an analytic tool for transparently and systematically judging whether available data support a mode of carcinogenic action hypothesized for an agent. It is not designed to give an absolute answer on sufficiency of the information as this will vary depending on the circumstance (IPCS, 2005). Amongst the strengths of the framework are its flexibility, general applicability to carcinogens acting by any mechanism and the ability to explore the impact of each key event on the carcinogenic response (IPCS, 2005). It is primarily based upon considerations for causality in epidemiologic investigations originally articulated by Hill (1965) but later modified by others and extended to experimental studies. The modified Hill criteria are useful in organizing thinking about aspects of causation, and they are consistent with the scientific method of developing hypotheses and testing those hypotheses experimentally. A key question is whether the data to support a mode of action meet the standards generally applied in experimental biology regarding inference of causation.

## **2.4 Components of a Mode of Action Analysis**

40. To perform a mode of action analysis, key biochemical, cellular, and molecular events need to be established, along with the temporal and dose-dependent concordance of each of the key events in the mode of action. The key events can be used to bridge species and dose for a given mode of action. The next step in the mode of action analysis is the assessment of biological plausibility for determining the relevance of the specified mode of action in an animal model for human cancer risk based on kinetic and dynamic parameters. (Boobis et al., 2006; Holsapple et al., 2005; Meek et al., 2003)

## **2.5 Postulated Mode of Action: Key Events**

41. The postulated mode of action is a biologically plausible hypothesized organization of the sequence of key events leading to the observed adverse health effect. The mode of action is supported by robust experimental observations and mechanistic data (IPCS, 2005 and Boobis et al., 2006). Key events are measurable events that are critical to the induction of tumors as hypothesized in the mode of action. To support an association, a body of experiments needs to define and measure an event consistently across dose and temporally relative to the other key events.

42. To evaluate whether an hypothesized or postulated mode of action is operative, an analysis starts with an outline of the scientific findings regarding the hypothesized key events leading to cancer. A weight of the evidence approach is used to evaluate the information to determine whether there is a causal relationship between these events and cancer formation. It is not generally expected that the complete sequence will be known at the molecular level. Instead, empirical observations made at different levels of biological organization (e.g., biochemical, cellular, physiological, tissue, organ, and system) are analyzed.

43. For each tumor site being evaluated, the mode of action analysis should begin with a description of all data relevant to the tumor of concern in general and specific for the chemical of interest. These data are described in relation to the key events that may be associated with an hypothesized mode of action and its sequence of key events. This can be followed by discussion of various aspects of the experimental support for hypothesized modes of action in animals and humans. For a more detailed discussion of the specific approach to mode of action analysis see Boobis et al., 2006 and 2008. See Appendix I. for some examples of mode(s) of action.

## **2.6 Experimental Support for the Postulated Mode of Action**

44. Experimental support addressing the strength, consistency and specificity of association, dose response concordance, temporal relationship and if the mode of action is biologically plausible all are necessary to establish a clearly articulated mode of action.

### ***1. Strength, Consistency and Specificity of Association***

45. A statistically significant association between events and tumor response observed in well conducted studies is generally supportive of causation. Consistent observations in a number of such studies with differing experimental designs increase that support, because different designs may reduce unknown biases. Studies showing recovery absence/reduction of carcinogenicity when the rate limiting event is blocked or diminished, are particularly useful tests of association. Conversely, if enhancement of rate limiting key events increases the tumor response, this evidence would also provide strong support for the postulated mode of action. Pertinent observations include tumor response and key events in the same cell type, sites of action logically related to key event(s), and results from multistage studies and from stop/recovery studies (Boobis et al., 2006). Specificity of the association without evidence of other modes of action also strengthens a causal conclusion. While these factors add strength to the mode of action, conversely, a lack of strength, consistency and specificity of an association tends to weaken the overall causal conclusions for a particular mode of action.

### ***2. Dose Response Concordance***

46. If a key event and tumor endpoints increase with dose such that the key events forecast the appearance of tumors at a later time or higher dose, the shape of the dose/response curve could be revealed and a causal association can be strengthened. Dose-response associations of the key event with other precursor events can add further strength.

### **3. Temporal Concordance**

47. If a key event is shown to be causally linked to tumorigenesis, it will precede tumor appearance. An event may also be observed contemporaneously or after tumor appearance; these observations may add to the strength of association but not to the temporal association. Pertinent observations include studies of varying duration observing the temporal sequence of events and development of tumors.

### **4. Biological Plausibility and Coherence**

48. The biological plausibility of any postulated mode of action in humans depends on a consideration of dose-effect and dose-response relationships (IPCS, 2005 and Boobis et al., 2006). The postulated mode of action and key events should be based on contemporaneous understanding of the biology of cancer. If the body of information under scrutiny is consistent with other chemical agents for which the hypothesized mode of action is accepted, the case is strengthened. Note: Because some modes of action can be anticipated to evoke effects other than cancer, the available toxicity database on noncancer effects can contribute to this evaluation.

## **2.7 Alternative Modes of Action**

49. All possible modes of action that could produce the adverse effect of concern should be considered and discussed. If there is evidence for more than one mode of action, each mode should receive a separate analysis. Furthermore, different modes of action can operate at different dose ranges; for example, an agent could act predominately at lower doses where cytotoxicity may not occur. Ultimately, however, information on all modes of action should be integrated to better understand how and when each mode acts, and which modes may be of primary interest for exposure levels relevant to humans.

## **2.8 Uncertainties, Inconsistencies and Data Gaps**

50. Uncertainties should be stated clearly, fully and explicitly. They should include those related to the biology of the toxicological response and those for the database on the specific chemical being evaluated. Any inconsistencies should be clearly noted and characterized with respect to the impact on the weight of evidence in support of the postulated mode action. Data gaps should also be identified and characterized. A clear statement should be included as to whether the identified data gaps are critical deficiencies in support of the postulated mode of action and recommendations should be provided for data needs to address those deficiencies (Boobis et al., 2008).

## **2.9 Conclusion of Postulated Mode of Action Analysis**

51. Conclusions about each postulated mode of action should address (1) whether the mode of action is supported in animals, (2) is relevant to humans, and (3) which populations or lifestages may be particularly susceptible. Special attention should be paid to whether these data suggest that tumors could arise after in-utero or childhood exposure. Because the cancer studies are usually performed with adult animals, conclusions about relevance during childhood generally rely on inference.

## **2.10 Relevance of rodent mode of action for humans**

52. "Relevance" of a potential mode of action is considered in the context of characterization of hazard and not at the level of risk. Anticipated levels of human exposure are not used to determine whether the postulated mode of action is operative in a particular population or lifestage, for example, in those with pre-existing disease (Boobis et al., 2006 and USEPA, 2005).

53. Other populations or lifestages may not be analogous to the test animals, in which case the question of relevance would be decided by inference. And although agent specific data would be most preferable, this review may also rely on general knowledge about the precursor events and characteristics of individuals susceptible to these key precursor events. Information suggesting quantitative differences between populations or lifestages should be flagged for consideration in the dose-response assessment, and a separate risk estimate should be quantified for susceptible populations or lifestage if data are available to describe a quantitative difference.

## 2.11 Human Relevance Framework

54. Considerable effort has been expended during the past several decades to evaluate the mode of action for specific chemicals causing cancer in rodents. However, the key question is the relevance of the particular postulated mode of action to human cancer risk assessment. A framework was developed through an ILSI/RSI working group sponsored by the U.S. EPA and Health Canada to address this issue and to provide direction in determining the relevance of rodent tumors to human health (Cohen et al., 2003; Meek et al., 2003; Cohen et al., 2004). The human relevance framework is not proscriptive and does not provide a check list of criteria; it is an analytical tool that describes a method using a decision tree logic to establish a relationship between early cellular events, the development of cancer, and its relevance to humans. Knowledge of key events and the identification of a mode of action provide a transparent and rational basis for human hazard and risk assessment.

55. The human relevance framework is based on three questions: (1) is the weight of evidence sufficient to establish the mode of action in animals, (2) are key events in the animal mode of action plausible in humans? and (3) taking into account kinetic and dynamic factors, are key events in the animal mode of action plausible in humans? [an alternative questions (2) and (3) based on the current IPCS framework (2) Can **human relevancy** of the MOA be reasonably **excluded** on the basis of fundamental, qualitative differences in key events between animals and humans? (3) Can **human relevancy** of the MOA be reasonably **excluded** on the basis of quantitative differences in either kinetic or dynamic factors between animals and humans?]. This is a more quantitative analysis which addresses the relevance of tumorigenicity to a level of exposure, and again relies on a concordance analysis between animal model and humans. This approach focuses not only on dose response but also on quantitative differences between species in fundamental biologic processes that can affect exposure dose associated with the adverse response.

56. Presentation in tabular form referred to as a concordance table can be particularly useful. The information in these tables should be relatively brief. In one column, the effect on humans for each key event evaluated and another column for the results in a different animal strain, species, or sex or for a different route of administration that does not result in toxicity. Factors may be identified that are not key events but can modulate key events and contribute to differences between species or individuals. Examples include genetic differences in pathways of metabolism, competing pathways of metabolism, and effects induced by concurrent pathology. While information for evaluating key events in humans may come from in vitro and in vivo studies on the chemical, basic information on anatomy, physiology, endocrinology, genetic disorders, human epidemiology, and other information that is known regarding the key events should be considered in this framework. (Boobis et al., 2008)

57. This human relevance framework is focused on hazard identification and evaluation. If the second and third questions are answered such that due to kinetic and/or dynamic differences there would not be a cancer hazard for humans then there is no cancer risk. (Holsapple et al., 2005). It is clearly acknowledged that departure from the default assumption of human relevance is a data rich determination, but if a conclusion is strongly supported by empirical data, exposure to chemicals producing the toxicity only via that mode of action would not pose a risk to humans, and therefore, no additional risk characterization for

this endpoint of carcinogenicity is further warranted (Boobis et al., 2008; EPA, 1991). An example of where a rodent mode of action was judged not to be relevant for humans was described for thiamethoxam-related mouse liver tumors. This determination was based on the quantitation of key metabolites in vivo and in vitro that showed mice, but not rats or humans, to be capable of generating sufficient amounts of the tumorigenic metabolites to initiate the hepatic toxicity necessary for tumor formation (Pastoor et al., 2005). See Appendix II for example of human relevance framework.

## 2.12 Hazard Characterization

58. Hazard characterization provides the overall weight of evidence summary of the assessment. It summarizes the conclusions about the agent's potential effects, whether they can be expected to depend qualitatively on the circumstances of exposure, and if there might be a susceptible subpopulation. It discusses the extent to which these conclusions are supported by data or are the result of default options invoked because the data are inconclusive. It explains how complex cases with differing results in different studies were resolved. The hazard characterization highlights the major issues addressed in the hazard assessment and discusses alternative interpretations of the data and the degree to which they are supportable scientifically.

59. When the conclusion is supported by mode of action information, the hazard characterization also provides a clear summary of the mode of action conclusions, including the completeness of the data, the strengths and limitations of the inferences made, the potential for other modes of action, and the implications of the mode of action for selecting viable approaches to the dose response assessment. The hazard characterization also discusses the extent to which mode of action information is available to address the potential for disproportionate risks in specific populations or lifestyles or the potential for enhanced risks on the basis of interactions with other agents or stressors.

## 2.13 For more information:

60. While this guidance is intended to provide an overall summary of the framework(s), for additional detail on the mode of action framework and human relevance framework, please refer to the IPSCS Framework for Analyzing the Relevance of a Cancer Mode of Action for Humans (Boobis, 2006) and USEPA Guidelines for Carcinogen Risk Assessment (USEPA, 2005)

([http://www.who.int/ipcs/methods/harmonization/areas/cancer\\_mode.pdf](http://www.who.int/ipcs/methods/harmonization/areas/cancer_mode.pdf))

### REFERENCES:

Boobis, AR., Doe, JE., Heinrich-Hirsch, B., Meek ME., Munn, S., Ruchirawat, M., Schlatter J, Seed J., Vickers, C. (2008) IPSCS Framework for Analyzing the Relevance of a Noncancer Mode of Action for Humans. *Critical Reviews in Toxicology*, 38:87-96.

Boobis, AR, Cohen, SM., Dellarco, V., Meek, ME., Vickers, C., Willcocks, d., Farland, W. (2006) IPSCS Framework for Analyzing the Relevance of a Cancer Mode of Action for Humans. *Critical Reviews in Toxicology*, 36:781-792. ([http://www.who.int/ipcs/methods/harmonization/areas/cancer\\_mode.pdf](http://www.who.int/ipcs/methods/harmonization/areas/cancer_mode.pdf))

EPA. Alpha-2u-globulin: association with chemically-induced renal toxicity and neoplasia in the male rat. EPA-625/3-91/019F. Washington, DC:U.S. Environmental Protection Agency, 1991.

Hanahan, D., and Weinberg, RA. (2000) The Hallmarks of Cancer. *Cell*, 100:57-70.

ENV/JM/TG(2009)X

Holsapple, MP, Pitot , HC, Cohen SM, Boobis, AR, Klaunig JE, Pastoor, T., Dellarco VL., and Dragan, YP. (2006) Mode of action in relevance of rodent liver tumors to human cancer risk. *Toxicol., Sci.* 89:51-56.

Pastoor T., Rose P., Lloyd S., Peffer R., and Green T. (2005) Case Study: Weight of Evidence Evaluation of the Human Health Relevance of Thiamethoxam-Related Mouse Liver Tumors. *Toxicological Sciences* 86(1), 56-60.

Rice, JM (2004) On the application of data on mode of action to carcinogenic risk assessment, *Toxicol. Sci.* 78, pp. 175–177.

USEPA (2005) Guidelines for Carcinogen Risk Assessment. U.S. Environmental Protection Agency, Risk Assessment Forum, Washington, DC March 2005. EAP/630/P-03/001F  
<http://cfpub.epa.gov/ncea/raf/recorddisplay.cfm?deid=116283>

**Appendix I. Examples of Animal Mode(s) of Action Framework**

Dellarco VL and Baetcke K. 1999. A Risk Assessment Perspective: Application of Mode of Action and Human Relevance Frameworks to the Analysis of Rodent Tumor data. *Toxicological Sciences* 2005 86(1):1-3; doi:10.1093/toxsci/kfi133

SAB, 1999, Minutes of the Chloroform Risk Assessment Review Subcommittee Meeting; Minutes of the Chloroform Risk Assessment Review Subcommittee Meeting, October 27-28, 1999, U.S. Environmental Protection Agency, Science Advisory Board. See website: <http://www.epa.gov/sab/chloro00.htm>

SAP, SAP Report No. 2000-05, FIFRA Scientific Advisory Panel Meeting, June 27- 29, 2000, Held at the Sheraton Crystal City Hotel, Arlington, Virginia, Atrazine :

Hazard and Dose-Response Assessment and

Characterization. See website: <http://www.epa.gov/oscpmont/sap/meetings/2000/june27/finalatrazine.pdf>

USEPA, 2005 Science Issue Paper: Mode of Carcinogenic Action for Cacodylic Acid

(Dimethylarsinic Acid, DMAV) and Recommendations for Dose Response

Extrapolation July 26, 2005, Health Effects Division, Office of Pesticide Programs, US Environmental Protection Program. See website;

[http://www.epa.gov/oppsrrd1/reregistration/cacodylic\\_acid/dma\\_moa.pdf](http://www.epa.gov/oppsrrd1/reregistration/cacodylic_acid/dma_moa.pdf)

SAB, 2007; Subject: Advisory on EPA's Assessments of Carcinogenic Effects of Organic and Inorganic Arsenic: A Report of the US EPA Science Advisory Board <http://www.epa.gov/sab/pdf/sab-07-008.pdf>

Manibusan.M, Odin M, DA Eastmond. 2007. Postulated Carbon Tetrachloride Mode of Action: A Review *Journal of Environmental Science and Health, Part C*, Volume 25, Issue 3 July 2007 , pages 185 - 209

McCarroll N, Keshava N, Cimino M, Chu M, Dearfield K, Keshava C, Kligerman A, Owen R, Protzel A, Putzrath R and R. Schoeny. 2008. An Evaluation of the Mode of Action Framework for Mutagenic Carcinogens Case Study: Cyclophosphamide. *Environmental and Molecular Mutagenesis* 49:117-131.

## **Appendix II. Example of Human Relevance Framework**

Pastoor T, Rose P, Lloyd S, Peffer R, and Green T. 2005. Case Study:Weight of Evidence Evaluation of the Human Health Relevance of Thiamethoxam-Related Mouse Liver Tumors. Toxicological Sciences 86(1), 56-60.

### **Abstract:**

Thiamethoxam was shown to increase the incidence of mouse liver tumors in an 18-month study; however, thiamethoxam was not hepatocarcinogenic in rats. Thiamethoxam is not genotoxic, and given the late life generation of mouse liver tumors, suggests a time related progression of key hepatic events that leads to the tumors. These key events were identified in a series of studies of up to 50 weeks that showed the time dependent evolution of relatively mild liver dysfunction within 10 weeks of dosing. Followed by frank signs of hepatotoxicity after 20 weeks leading to cellular attrition and regenerative hyperplasia. A metabolite CGA330050 was identified as generating the mild hepatic toxicity, and another metabolite, CGA265307, exacerbated the initial toxicity by inhibiting inducible nitric oxide synthase. This combination of metabolite generated hepatotoxicity and increase in cell replication rates is postulated as the mode of action for thiamethoxam and relevance of these mouse specific tumors to human health was assessed by using the framework and decision logic developed by ILSI/RSI. The postulated mode of action was tested against the Hill criteria and found to fulfill the comprehensive requirements of strength, consistency, specificity, temporality, dose response, and the collective criteria of being a plausible mode of action that fits with known and similar modes of action. Whereas the postulated mode of action could theoretically operate in human liver, quantitation of key metabolites in vivo and in vitro showed that mice, but not rats or humans, generate sufficient amounts of these metabolites to initiate the hepatic toxicity and consequent tumors. Indeed rats fed 3000 ppm thiamethoxam for a lifetime did not develop hepatotoxicity or tumors. In conclusion, the coherence and extent of the database clearly demonstrates the mode of action for mouse liver tumorigenesis and also allows for the conclusion that thiamethoxam does not pose a carcinogenic risk to humans.

### 3. STUDY DESIGN

#### 3.1 Dose Selection

##### 3.1.1 Introduction

61. The purpose of a long-term toxicity or carcinogenicity study is the detection of biological evidence of any toxic or carcinogenic potential of the substance being investigated. Protocols should therefore maximise the sensitivity of the test without significantly altering the accuracy and interpretability of the biological data obtained. The dose regimen has an extremely important bearing on these two critical elements. Since one of the objectives is determination of the dose–response relationship in respect of any observed effects, the OECD Test Guidelines 451, Carcinogenicity Studies, TG 452, Chronic Toxicity Studies and TG 453, Combined Chronic Toxicity/Carcinogenicity Studies normally require at least three dose levels, as well as controls.

62. OECD Test Guidelines 451, 452 and 453 outline general principles for dose selection in a carcinogenicity/chronic bioassay. Provision of in depth guidance and a strategy for dose selection is however beyond the scope of the test guideline texts. This section of the guidance document is designed to underpin and expand the principles of dose selection for chronic toxicity and carcinogenicity studies outlined in the test guidelines.

63. As noted in section 1.2 on scope of application of the guidance, these principles of dose selection are generally applicable to a wide range of chemicals, whatever their field of application, e.g. pesticides, industrial chemicals and pharmaceuticals. However, although this chapter provides a number of references to specific requirements for dose selection for pharmaceuticals, the principles applied in studies on pharmaceuticals may differ from that for other agents (Rhomberg et al., 2007; ICH, 2008). More information is generally available on the pharmacodynamic effects of pharmaceuticals, including the results of controlled clinical studies. The intended systemic human exposure is known and detailed pharmacokinetic studies enable valid comparisons to be made between the systemic exposures in rodents at the chosen dose levels and those in humans under therapeutic administration of the drug, as measured by the comparative areas under the curve (AUC) of blood concentrations over time (Rhomberg et al., 2007). Users of the guidance should therefore consult the [Guideline S1C on dose selection for carcinogenicity studies of pharmaceuticals for specific information on testing of such chemicals \(ICH, 2008\)](#)

64. The core dose selection strategy is dependent on the primary objective or objectives of the study (ILSI, 1997; Rhomberg et al. 2007), namely:

- identification of the hazardous properties of a chemical,
- characterisation of the dose:response relationship,
- identification of a threshold or Benchmark Dose departure point,
- the provision of information on the health effects at human exposure levels, and/or
- provision of data to test hypotheses regarding mode of action (MoA).

65. In selecting appropriate dose levels for a carcinogenicity/chronic toxicity bioassay, a balance has to be achieved between hazard screening on the one hand and characterisation of low-dose responses and their relevance on the other. This is particularly relevant in the situation where a combined chronic toxicity and carcinogenicity study (TG 453) is to be carried out. While this chapter of the guidance goes on to discuss different study designs dependent on the primary objective of the study, it is essential to recognise that for animal welfare and economic reasons the chronic toxicity/carcinogenicity potential of a chemical will normally only, in the majority of cases, be investigated in a single bioassay in one species (see chapter 3.3 for guidance on choice of species and testing in more than one species). The design of this single study must therefore be carefully planned in order to maximise the information to be obtained from it. The information provided on differing objectives and different study designs in the guidance should not be construed as an indication that more than one bioassay should be carried out on a particular chemical.

66. As already outlined in the Introduction to Chapter 1 of this guidance, general principles and guidance on dose selection for chronic toxicity and carcinogenicity studies in rodents are provided in two reports of the International Life Sciences Institute (ILSI). An initial 1997 report, entitled *Principles for the Selection of Doses in Chronic Rodent Bioassays* (ILSI, 1997), presented common views on the selection of doses for chronic carcinogenicity and toxicity studies while a second ILSI working group publication in 2007, entitled *Issues in the Design and Interpretation of Chronic Toxicity and Carcinogenicity Studies in Rodents: Approaches to Dose Selection* (Rhomberg et al. 2007) provides additional discussion of the factors that influence dose selection in chronic bioassays (Rhomberg et al., 2007). The latter publication incorporates concepts included in other documents prepared by national and international organizations (OECD, ECETOC, NTP and USEPA), and places emphasis on the influence of the objectives of a chronic bioassay on dose selection, as summarised in section 3.1.3 of this guidance. Users of this guidance document are recommended to consult these publications for more information on the factors influencing dose selection.

67. OECD gratefully acknowledges the fact that ILSI made this information available to the OECD before publication in the open literature and thus enabled the development of this Guidance Document. The current chapter of the Guidance Document was developed on the basis of full consideration of both the views expressed in the ILSI reports and the views of OECD experts developed at a workshop on the revision of OECD Test Guidelines 451, 452 and 453 held in Washington on 26-28 February, 2008.

68. The following sections provide guidance on (a) the general principles for dose selection identified in the Test Guidelines 451, 452, 453, (b) the influence of the objectives of a chronic bioassay on dose selection. It should be noted, however, as already stated in Chapter 1 and as discussed in greater detail by Rhomberg et al. in section 2.2 of their publication, that the basic principles for the conduct of chronic toxicity and carcinogenicity studies will differ, given that the endpoints are different. However, given the drive to reduce the number of animals used for welfare reasons and the cost of carcinogenicity bioassays, there is a need to maximise the results to assess non-cancer effects that may arise during the study, as these may be critical to the interpretation of any carcinogenic effects. The possibilities for doing so are maximised in the TG 453, Combined Chronic Toxicity/Carcinogenicity Study.

69. While the guidance provided in this Chapter can be taken as generally applicable to dose selection for either a chronic toxicity or a carcinogenicity bioassay, or a combined chronic and carcinogenicity study, users of the guidance should be mindful of the primary objectives of the study in establishing the optimum study design in terms of dose selection. For example, in a chronic toxicity bioassay, the doses selected will include at least one dose high enough to show toxicity, at least one dose low enough to show lack of toxicity, and one or a few doses in between to help characterize the shape of the curve near the point where the threshold appears to lie (Rhomberg et al., 2007). These dose placement concerns differ from those in the typical carcinogenicity bioassay.

### 3.1.2 Objectives of a Chronic Bioassay

70. The general principles for dose selection laid down in the test guidelines are summarised as follows:

- Dose levels will generally be based on the results of shorter-term repeated dose or range finding studies and should take into account any existing toxicological and toxicokinetic data available for the test substance or related materials.
- the highest dose level should be chosen to identify the principal target organs and toxic effects while avoiding suffering, severe toxicity, morbidity, or death;
- dependent on the objectives of the study; a top dose lower than the dose providing evidence of toxicity may be chosen, e.g. if a dose elicits an adverse effect of concern that nonetheless has little impact on lifespan or body weight. The top dose should not exceed 1000 mg/kg body weight/day, except in the case of pharmaceuticals, discussed further below.
- Dose levels and dose level spacing may be selected to establish a dose:response and, depending on the mode of action of the test substance, a NOAEL or other intended outcome of the study, e.g. a BMD at the lowest dose level.
- Factors that should be considered in the placement of lower doses include the expected slope of the dose–response curve, the doses at which important changes may occur in metabolism or mode of toxic action, where a threshold is expected, or where a point of departure for low-dose extrapolation is expected.
- points to be considered in dose selection include:
  - known or suspected nonlinearities or inflection points in the dose–response;
  - pharmacokinetics, and dose ranges where metabolic induction, saturation, or nonlinearity between external and internal doses does or does not occur;
  - precursor lesions, markers of effect, or indicators of the operation of key underlying biological processes;
  - key (or suspected) aspects of mode of action, such as doses at which cytotoxicity begins to arise, hormone levels are perturbed, homeostatic mechanisms are overwhelmed, etc.;
  - regions of the dose–response curve where particularly robust estimation is required, e.g., in the range of the anticipated BMD or a suspected threshold;
  - consideration of anticipated human exposure levels

71. These principles for dose selection are broadly similar to the key principles for dose selection outlined in the ILSI publications (ILSI, 1997; Rhomberg et al., 2007), as listed in full in the Appendix to this guidance document. They are further discussed in the following sections, under the headings of:

- key information for the selection of doses in chronic toxicity and carcinogenicity studies,
- selection of the highest dose to be used,
- dose level spacing.

#### 3.1.2.1 Key information for the selection of doses in chronic toxicity and carcinogenicity studies

72. As stated in the first bullet point in section 3.1.2 above “Dose levels will generally be based on the results of shorter-term repeated dose or range finding studies and should take into account any existing toxicological and toxicokinetic data available for the test substance or related materials.” Ultimately the robustness of a carcinogenicity or chronic toxicity study, in particularly the former, is dependent on a demonstration that the dose levels selected in the study have been adequate to demonstrate an effect or effects of the chemical, without producing either false negative results (because the doses selected were too low) or false positive results (because metabolic/homeostatic mechanisms are overwhelmed, etc) which cannot be used in assessing risk in humans.

73. The 2007 ILSI report (Rhomberg et al., 2007) provides a list of parameters and changes relative to controls to be considered when evaluating the acceptability of dose levels included in a carcinogenicity study, in particular in reaching a decision as to whether an adequate high dose or maximum tolerated dose (MTD) has been used (Appendix 2 of Rhomberg et al., 2007). These included clinical signs, reductions in body weight gain, organ weight increases, changes in haematological and clinical chemistry parameters, endocrinology disturbances and pathological changes, particularly in the major organs such as liver and kidney. The ILSI report built on two earlier publications, from ECETOC (ECETOC, 1996) and that of Smith and colleagues (Smith et al. 2002).

74. The data provided by shorter-term repeated dose or range finding studies, including 28-day or 90-day studies, are critical in selecting the dose levels for a longer-term chronic toxicity or carcinogenicity study. The dose levels used in such studies and the NOAELs established can be used as a starting point for dose selection, both in relation to the highest dose level to be chosen in the study and possibly (but not necessarily) to the lower dose levels. However, in addition, the parameters that may be evaluated for assessing the adequacy of the dose levels in a completed carcinogenicity study as proposed by ILSI and others (Rhomberg et al., 2007; ECETOC, 1996; Smith et al. 2002) can also be applied to the results of these shorter-term to provide criteria for dose selection for a longer-term study (Rhomberg et al., 2007). Considerations that should be taken into account in determining whether similar, lower or higher dose levels than those used in a short-term study should be selected for a chronic toxicity or carcinogenicity study include (Rhomberg et al., 2007):

- whether the effect is an adaptive response (e.g., liver hypertrophy in the absence of any other evidence of hepatotoxicity);
- potential of the toxic effect(s) observed in prechronic studies to progress to neoplasia. A dose that induces a marked effect in a prechronic study should not be excluded from a carcinogenicity study if the effect or effects can reasonably be anticipated to be a precursor event in the development of neoplasia (e.g., a key event for the mode of action of the chemical). However, care should be taken that selection of a dose level that induces such effects will not result in excessive toxicity in the carcinogenicity study;
- the potential that an effect may limit the sensitivity of the chronic/carcinogenicity study (e.g., haemolytic anemia that may lead to an increase in mortality or otherwise compromise the health of the animal);
- the duration of the short-term study (e.g., 90-day subchronic study, 28-day subacute study, two-generation reproduction study) and the potential for a toxic effect to progress in severity (e.g., progression from focal to multifocal necrosis);
- evidence that an observed toxic effect in a prechronic study is transitory (e.g., an increase in TSH levels). The dose that induces a toxic effect that is transitory may be selected as the highest dose in a chronic toxicity or carcinogenicity study, as it is not expected to progress in severity other than to neoplasia; and
- use of gavage for test chemical administration in a prechronic study. A dose that induces overt toxicity in a gavage study may be tolerated if a dietary route of administration is selected for a carcinogenicity study because of the differences in toxicokinetics and toxicodynamics between the two methods of administration.

75. Additional evidence on the extent to which dose levels should be increased or decreased in a long-term study relative to a short-term study or studies may be provided by dose–response data from the latter studies. For example, a marked reduction in dose levels would be warranted if results from short-term studies show that a minor increase in dose is associated with a pronounced increase in severity or incidence of a lesion (i.e., a steep dose–response). It is recommended that the use of an arbitrary factor (e.g., one-tenth the highest dose tested in a short-term study that

induced a severe toxic effect) should be avoided when selecting the high dose (or mid and low dose levels) for a proposed carcinogenicity study.

76. Toxicokinetic data should always be taken into account when selecting dose levels for a chronic toxicity or carcinogenicity study, although such data may not be readily available for all chemicals, as they are not required under all regulatory schemes. Many toxicokinetic processes influencing absorption, distribution, elimination and metabolic activation or detoxication may become saturated at higher doses, resulting in systemic exposures to parent compound or metabolites that would not be expected in the real life human exposures for which risk assessments are needed. The effect of repeated exposures on the pattern of absorption, metabolism, detoxification, and clearance of a compound will provide information on the internal dose achieved during chronic exposure under conditions of the bioassay. The importance of having data on toxicokinetics in reaching a decision on the optimum design for a chronic toxicity or carcinogenicity study is stressed in this guidance, and the use of such data is discussed in more detail in chapter 3.4 of this guidance document.

77. Physiologically-based toxicokinetic (PBTK) modelling is also a valuable tool for defining doses where non-linear toxicokinetics may occur, thus allowing this to be considered in selecting the highest and other dose levels in the study. The use of PBTK modeling is also explored in more detail in chapter 3.4. Finally, specific mechanistic studies may provide useful information regarding target tissues affected by the chemical and the doses associated with effects on key events, and should be taken into account when selecting doses for a chronic toxicity or carcinogenicity study.

78. Additional considerations in selecting dose levels for chronic toxicity or carcinogenicity studies arise as a result of practical constraints such as the physicochemical characteristics of the substance to be tested (e.g., solubility, vapour pressure), palatability of the compound in food or drinking water, and other factors such as the potential for the substance to cause adverse effects such as irritancy at the site of administration (8). Further guidance is provided on these aspects is provided in Chapter 3.5 of this guidance and also in the ILSI publications (ILSI, 1996; Rhomberg et al., 2007).

### 3.1.2.2 *Selection of the highest dose*

79. Dose selection should be based on the findings of subchronic or other range-finding studies. The highest dose level to be used in a chronic toxicity or carcinogenicity study needs to be carefully considered and the reasons for the final choice clearly defined. Ideally, the dose levels selected will maximise the detection of dose–response relationships and facilitate the extrapolation of these to potential hazards for other species, including humans. The largest administered dose should **not** compromise the biological interpretability of the observed responses.

80. The selection of the highest dose level to be used in a chronic toxicity or carcinogenicity study has long been a matter of controversy. Bullet point (2) in section 3.1.2, reproduced from the Test Guidelines, states that “*the highest dose level should be chosen to identify the principal target organs and toxic effects while avoiding suffering, severe toxicity, morbidity, or death*”

81. At the time when long-term (chronic) animal bioassays began to be routinely used to assess the qualitative potential of a chemical to cause chronic toxicity and cancer, the emphasis was on testing at high levels in order to maximize the potential of such studies to detect effects. The concept of the Maximum Tolerable Dose (MTD), often defined as the highest dose to produce toxic effects without causing death and to decrease body weight by no more than 10% relative to controls (16) became well established. More recently use of the MTD has been challenged, with increasing emphasis being placed by government regulatory bodies on formal risk assessment procedures and relevance for humans. While the concept of the MTD is still applied to some extent in selection of the highest dose to be used in a chronic toxicity or carcinogenicity study, there is a lack of consensus on how this should be defined (Rhomberg et al., 2007).

82. While some regulatory bodies or organizations interpret an adequate high dose to be a minimally toxic dose, others emphasize the need to select a dose level that is a maximally tolerated dose (i.e., more severe toxicity should be demonstrated). Thus, because of differences in views regarding the severity of toxic effects that are interpreted as providing evidence that an adequate high dose has been attained or exceeded, a completed carcinogenicity bioassay may be considered to be acceptable by one organization but not by another. Appendix 2 of Rhomberg et al. (2007) provides detailed guidance on criteria that can be applied in order to assess the acceptability of the high dose level or MTD (Rhomberg et al., 2007).

83. For non-genotoxic substances where thresholds may exist and carcinogenicity may result from alterations in normal physiology, linear extrapolations from high dose effects have been questioned. This has led to the concern that exposures in rodents greatly in excess of the intended human exposures may not be relevant to human risk; because they so greatly alter the physiology of the test species, the findings may not reflect what would occur following human exposure (ICH, 2008).

84. These considerations must be kept in mind when selecting the highest dose for such a study. Many carcinogenicity studies can be challenged on the basis of selection of a top dose that is too high, usually as a result of MTD considerations, often combined with an over-conservative approach to the dose selection of the other dose levels. This results in data that are difficult to interpret and cannot be used for regulatory purposes. Where the term MTD is used in the rest of this section, it is used to describe “the highest dose level to be used in a chronic toxicity or carcinogenicity study”. However, the interpretation of this in terms of the dose selection and the ultimate acceptability of the study for regulatory purposes must always be considered.

85. If the main objective of the study is to identify a cancer hazard (see also section 3.1.3.1), there is broad acceptance that the highest dose should not cause excessive toxicity (as indicated by substantial (20% or more) body weight loss and/or histopathological evidence of target organ toxicity) and should not be anticipated to shorten the test animal's life expectancy for reasons other than the development of tumors. Excessive toxicity at the top dose level (or any other dose level) may compromise the usefulness of the study and/or quality of data generated. Criteria that have evolved for the selection of an adequate high dose level, as already mentioned in 3.1.2.1, include: (in particular) toxicokinetics; saturation of absorption; results of previous repeated dose toxicity studies; the Mode of Action (MoA) and the maximum tolerable dose (MTD).

86. The general approach to selecting the maximum dose in a study should be based on a tiered approach, taking MoA into consideration. For compounds that are (or might be) genotoxic, conventional considerations of MTD given above would apply. For compounds that are not genotoxic, the maximum dose should be informed by considerations of MoA. For a given compound, if the Margin of Exposure (MOE) based on long term studies of hepatic toxicity is acceptable, then this MOE would also be protective against any carcinogenic effect. This is because non-genotoxic carcinogens produce cancer by perturbing normal physiology or biochemistry. The chronic assay should be designed to identify and characterise these perturbations and not necessarily cancer *per se*.

87. Operationally an adequate high dose (MTD) to be used in a chronic bioassay ideally produces some minimal signs of toxicity such as slight depression of body weight gain (not more than 10%), normally without causing tissue necrosis or metabolic saturation and without substantially altering normal life span due to effects other than tumors.

88. Most national and international organizations consider a decrement in body weight gain approaching 10% over the lifetime of the animals as definitive evidence that a high dose has been selected for or achieved in a carcinogenicity study (ILSI, 1997). However it is recognized that metabolic saturation or tissue necrosis is a key event in the MOA of a number of carcinogens (see chapter 2), and it may be necessary to deviate from these principles regarding the top dose level to be used in order to fully explore the carcinogenic potential and MOA of the chemical in question. It is also recognized that a top dose lower than the dose providing evidence of toxicity may be chosen, e.g. if a dose elicits an adverse effect of concern that nonetheless has little impact on lifespan or body weight, such as immunotoxicity (Rhomberg et al., 2007).

89. Nutritional effects, physiological factors, physical-chemical factors and compound bioavailability can influence selection of the highest dose level to be used in a chronic bioassay. For nutritional and possibly other physiological reasons a maximum level is imposed – commonly 5% concentration in the diet.

90. The Limit Dose for pesticides and other industrial chemicals in all cases is 1000 mg/kg bw/day. . However, according to ICH S1C(R2), the limit dose for carcinogenicity studies for pharmaceuticals is 1500 mg/kg bodyweight per day, provided that this dose results in systemic exposure that is 10 times the expected human systemic exposure. For some pharmaceuticals, exposure after a nominal dose of 1000 mg/kg bodyweight per day may not exceed the anticipated human exposure. As indicated in section 3.1.2.1, palatability of a compound in either feed or water can also perturb physiological homeostasis or nutritional status. A compound's solubility limit or vapor pressure may constrain selection of the top dose level. Irritation at the site of compound deposition may constrain dose or otherwise confound cross species extrapolation. Inhalation of doses that overwhelm pulmonary clearance may lead to tissue responses that are specific to the species being tested.

### 3.1.2.3 Dose level spacing

91. Selection of dose intervals is influenced by the study objectives (see section 3.1.3) and the available information. Dose levels and dose level spacing may be selected to establish a dose:response and, depending on the mode of action of the test substance, e.g., for non-genotoxic carcinogens and in the case of a chronic toxicity study, a NOAEL or other intended outcome of

the study, e.g. a BMD at the lowest dose level. The dose level spacing does not need to be regular. The increasing emphasis on consideration where the lower dose levels used in the study are placed, and the number of such dose levels, reflects the changing purposes of lifetime bioassays.

92. If the primary purpose is characterization of hazard, whether this is chronic toxicity or carcinogenicity, the focus of dose selection should be on maximizing the power of the study and on the top doses tested. As the risk assessment process becomes increasingly concerned with characterization of human risk, there has been a corresponding need to characterize whether and how high-dose effects extend to responses at lower exposure levels as well, with a consequent interest in how the lower dose levels are placed in bioassays (Rhomberg et al., 2007)

93. As outlined in the Test Guidelines and as discussed in detail in the ILSI publications (ILSI, 1997; Rhomberg et al., 2007): dose selection and dose level spacing need to be based on the following considerations:

- known or suspected nonlinearities or inflection points in the dose–response;
- pharmacokinetics, and dose ranges where metabolic induction, saturation, or nonlinearity between external and internal doses does or does not occur;
- precursor lesions, markers of effect, or indicators of the operation of key underlying biological processes;
- key (or suspected) aspects of mode of action, such as doses at which cytotoxicity begins to arise, hormone levels are perturbed, homeostatic mechanisms are overwhelmed, etc.;
- regions of the dose–response curve where particularly robust estimation is required, e.g., in the neighborhood of the anticipated BMD;
- a suspected threshold; *and*
- anticipated human exposure levels.

94. Lower dose levels should not automatically be set as fractions of the highest dose used in the study, e.g. 50% of the top dose, 25% of the top dose, etc. (ILSI, 1997; Rhomberg et al., 2007). Rather, dose levels should be selected to reflect the purposes of the study, and they should use available knowledge on how dose-dependent biological and impacted physiological factors may affect study outcomes. The Test Guidelines indicate that “The dose level spacing selected will depend on the characteristics of the test substance, and cannot be prescribed in this Guideline, but two to four fold intervals are frequently optimal for setting the descending dose levels and addition of a fourth test group is often preferable to using very large intervals (e.g., more than a factor of about 6-10) between dosages. In general the use of factors greater than 10 should be avoided, and must be justified if used.

#### 3.1.2.4 *Relevance of effects at high dose levels*

95. A design often applied uses a mid dose that is half of the top dose, or the geometric mean of the low and high dose. It may be possible to place adjacent dose levels somewhat above and below the levels at which a key transition in underlying biological actions is believed to lie,

thereby revealing its influence on response. Transitions need not be sharp; typically, there are ranges of doses over which an underlying biological factor, such as metabolic saturation or cytotoxicity, comes increasingly into play. The aim is to place doses so that the role of such underlying phenomena can be revealed. The issue of where to place the lowest dose should receive comparable attention to the placement of the top dose. If the lowest dose is too low, it may be insufficiently powerful and therefore uninformative; if too high, it may lose opportunities to characterize effects as near as possible to environmental exposure levels.

### ***3.1.3 Objectives of a chronic bioassay and their influence on dose selection***

96. As already outlined in section 3.1.2, the ILSI publications (ILSI, 1997; Rhomberg et al., 2007) provide practical guidance on factors that influence dose selection in chronic bioassays, with particular emphasis on how the varying objectives of a chronic toxicity/carcinogenicity bioassay influence dose level selection. This part of the guidance goes on to discuss different study designs dependent on the primary objective of the study. However, as already noted in the Introduction to this chapter it is essential to recognise that for animal welfare and economic reasons the chronic toxicity/carcinogenicity of a chemical will normally only, in the majority of cases, be investigated in a single bioassay in one species (see chapter 3.3 for guidance on choice of species and testing in more than one species). The design of this single study must therefore be carefully planned in order to maximise the information to be obtained from it. The information provided on differing objectives and different study designs in the guidance should not be construed as an indication that more than one bioassay will be carried out on a chemical under test. Rather, it is intended to provide specific guidance in the case where a particular objective has been identified for the study.

97. The ILSI publications (ILSI, 1997; Rhomberg et al., 2007) distinguish six possible objectives of a chronic bioassay and provides, for each of them separately, the theoretical best approach for dose selection:

- Screening chemicals to identify carcinogens or those causing other toxic effects)
- Characterizing the dose–response curve in the observable range
- Characterizing the dose–response curve to facilitate low-dose (linear) extrapolation
- Defining a threshold or BMD departure point
- Providing data on health effects at human exposure levels
- Providing data to test hypotheses regarding mode of action

98. These six objectives are discussed below.

#### ***3.1.3.1 Screening Chemicals to Identify Carcinogens (or other toxic effects)***

99. This may be the primary objective of the OECD Test Guideline 451 on Carcinogenicity Studies, and also the TG 453 Combined Chronic Toxicity/Carcinogenicity Studies, for which the core minimum study design comprises at least four groups (one control and at least three

treatment groups), each of which is exposed to different concentrations of the test substance. Assuming this is the primary objective of the study, a high dose close to/at a maximally-tolerated/minimally-toxic dose (e.g. MTD) should be chosen (see section 3.1.2 on consideration of the highest dose level in the study and selection of dose level spacing). The next dose may be set relatively close to (e.g. half of) the selected top dose, with other dose levels being set proportionately. This will ensure that the power and sensitivity of the assay is maximized and that at least one dose will not have a carcinogenic or other effect. This approach minimizes the chance of a false negative (failing to detect an effect that actually exists) at some increased risk of a false positive (finding a high-dose effect that is an artifact of excessively high doses and is not relevant to the dose range of interest).

100. The number of dose groups and number of animals in each dose group and the spacing of these doses must be chosen to meet as far as possible the study objective. Depending on the aim of the study, it may be possible to increase the statistical power of the key estimates by allocating animals more than 50 animals to the low dose groups; e.g., to estimate the carcinogenic potential at low dose levels. However, as indicated in TG 451, it should be recognized that a moderate increase in group size will provide relatively little increase in statistical power of the study. Further information on statistical design of the study and choice of dose levels to maximise statistical power is provided in section 3.7 of this guidance document. The use of two low doses may not always be necessary, since the lower dose serves (under this objective) as a hedge against discovery after-the-fact that the top dose unacceptably exceeds a maximally-tolerated dose (Rhomberg et al., 2007).

101. The advantage of this approach is that the results are relatively straightforward to interpret and widely acceptable. The data generated should provide a clear answer to the question of whether a chemical can cause a particular toxic response. A "verified and validated" negative outcome provides evidence that the compound lacks the ability to cause toxic effects in a particular animal bioassay, since lower doses would be even less powerful in detecting an effect. However, a limitation of the approach is that it will provide minimal data regarding the shape of the dose–response curve in the low-dose range.

102. Although the dose–response data gained are slightly superior to those that might be obtained with a single dose level, the support for low-dose extrapolation is minimal. The data will provide little information about possible nonlinearities in the dose–response curve or the existence of a threshold. The relevance of the high dose level recommended to be used in the study (close to a maximally-tolerated dose) to potential human exposures can also be debated. Mechanistic information gleaned from this type of study may be irrelevant. If the top dose level is set lower, to ensure relevance, the power to detect effects may be compromised. In short, positives may be difficult to interpret vis-à-vis low exposure levels, because they may reflect a high-dose-only phenomenon.

### 3.1.3.2 *Characterizing the Dose–Response Curve in the Observable Range*

103. If this is the objective of the study, which may be the case in particular for TG 452, but also may be an objective for TG 451 and 453, dose levels should be spread out to ensure determination a reliable dose-response curve. A minimum of three dose levels should be used; more doses may be appropriate if nonlinearities are expected and if the effects of interest occur

frequently (i.e., if power to detect rare events is not critical). Doses should be dictated by the possible dose-response that could appear, and the dose range expected effects, (i.e., observable range.). The top dose level should be set at or close to a maximally-tolerated/minimally toxic dose.

104. Use of a maximally-tolerated dose as the highest dose level in such a study may have little relevance except as an indicator of the ability of the assay to find adverse effects. Indeed, depending on the endpoint of interest, the top dose level may be well below the conventional MTD. Animal numbers should be distributed fairly evenly across the range of dose levels if very little is known about the nature of the response prior to the experiment. If prior evidence allows, it may be possible to optimize the design in terms of the location of the dose levels and the allocation of animals to them.

105. This approach will provide information regarding changes in the toxicity of the chemical with increasing dose, providing some possible insights as to the mode of action. It allows for extrapolation to health effects at meaningful human exposures, and still retains some limited ability for hazard identification. However, the assay will have less power to detect effects compared with the approach outlined for the objective outlined in section 3.1.3.1 unless additional animals are used. For purposes of hazard identification, the likelihood of a false-negative result, even at an MTD, will be increased. The utility in defining the dose-response curve will depend on how well the dose placement anticipates the interesting and informative parts of the curve; if the chosen dose levels “miss” these points, the results may not be informative.

### *3.1.3.3 Characterizing the Dose-Response Curve to Facilitate Low-Dose (Linear) Extrapolation*

106. If low-dose extrapolation is the primary objective of the study, as may be the situation for some carcinogenicity studies, and if the shape of the dose-response curve at relatively high dose levels is considered marginally relevant (as is the case with the point-of-departure-and-linear-extrapolation procedure), all of the dose levels chosen should be placed at the low end of the dose-response range, so that even if they show no significant dose-response, the fitted curve is forced to be as low and flat as possible. The lowest linear “upper-bound” extrapolation consistent with any actual effects that may exist is achieved if a dose level anchors the observed dose-response relationship at just under the dose where observable effects begin to occur. If nonlinearities are expected, the dose levels should be placed below the inflection point, if it exists. Dose location and animal number should be selected in order to minimize the standard error of low-dose estimates.

107. This approach would best facilitate extrapolation to health effects at meaningful human exposures and would provide a bounding estimate of the dose-response slope. When low-dose extrapolations are linear or upper bounds, this approach minimizes the tendency to overestimate low-dose risks. However the study will provide no data about the shape of the dose-response curve in the observable range and will be of little or no value for hazard identification. Indeed, by avoiding high dose levels, actual positive responses might not be detected. The approach is most useful when the substance’s carcinogenicity has already been established or can be predicted, and the question is solely one of obtaining a linear low-dose extrapolation. If the true dose-response curve is nonlinear, the lack of information about shape will make it difficult to choose a low-dose

extrapolation approach suited to non-linear responses (such as the margin-of-exposure approach) over a low-dose linear extrapolation.

#### *3.1.3.4 Defining a Threshold or BMD Departure Point*

108. Definition of a threshold or BMD departure point can be a primary objective for both carcinogenicity and chronic toxicity studies. Information essential for defining a threshold or BMD departure point in such studies will be derived from an initial 28-day or 90-day toxicity study employing a number of dose levels. However, it will be also necessary to apply a BMD approach in these preliminary studies, in order to accurately determine dose levels for the long-term study.

109. To achieve the objective, the number of dose groups may need to be higher than the identified core minimum study design identified, which normally comprises four dose groups (a control and three treatment groups). As indicated under section 3.1.3.1. and as discussed in more detail in chapter 3.7 of this guidance document on statistical design of the study and choice of dose levels to maximise statistical power, for a study with this objective it may be possible to increase the statistical power of the key estimates by differentially allocating animals to the various dose levels, e.g., to estimate the carcinogenic potential at low doses it may be necessary use more than 50 animals in the low dose groups. The increased number of dose groups may also be balanced by using fewer animals per group throughout the study design. However, this approach is not advisable for initial or one-off studies, but should only be used if there is information from other sources to support differential allocation of animals or focusing the dosing in a narrow range.

110. The dose range used should be wide enough to include a high dose level, associated with a clearly identified effect, and a low dose level associated with a smaller but detectable effect (preferably lower than the BMD). It should be noted that modelling using the benchmark dose approach will not be possible unless dose levels are high enough to produce some effect at one or (preferably) more doses. Dose levels and animal numbers should be concentrated in the region of the anticipated NOAEL or low response.

111. As indicated above, a minimum of three dose levels are likely to be required, with the high dose level used mainly as a check on the ability of the assay to detect adverse effects. Dose levels in the middle portion of the dose–response curve may be reduced, allowing the additional dose groups in the study to be focused primarily at the lower part of the dose:response curve. The requirements of paragraph 24 of TG 451 and TG 452 and paragraph 26 of TG 453 should be noted in relation to a study designed to meet this objective, that two to four fold intervals are frequently optimal for setting the descending dose levels and addition of a fourth test group is often preferable to using very large intervals (e.g., more than a factor of about 6-10) between dosages. As indicated in the Test Guidelines, in general the use of factors greater than 10 should be avoided, and should be justified if used.

112. The information provided by this approach should enable the definition of a NOAEL or threshold of toxicity and because a dose-response curve is generated to determine the BMD, it may also provide information on potency, helping to answer the question “How much is toxic”? The data generated in such a study are more likely to be relevant to human exposure levels than

e.g. those obtained from a study designed to screen chemicals to identify carcinogens (or other toxic effects (section 3.1.3.1)).

113. A toxic effect at the high dose level may be sufficient for hazard identification, however the approach will be of limited value should no substance-related changes be identified at the highest dose level. Limited information may be obtained regarding the shape of the dose–response curve, particularly if non-linearity is seen in the middle of the dose range. The power of the assay at lower dose levels will also be limited if the incidence of the responses of interest is low (e.g. rare tumours). If the doses tested are too low, the incidence of responses of interest will not be markedly different from those among controls. In general, the utility of the approach depends on how well the dose range of interest is anticipated.

#### *3.1.3.5 Providing Data on Health Effects at Human Exposure Levels*

114. In this approach, dose levels used in the study are normally based on a defined level of the target population’s exposure of interest and multiples of that exposure (e.g., 10 times or 50 times higher). It should be noted that such an approach may be useful in the case of testing of pharmaceuticals but is not likely to be useful in crop protection or commodity chemicals, given the uncertainties regarding exposure levels in scenarios where these are used.

115. If pharmacokinetic data are available, dose levels based on internalized doses (e.g., Area Under Curve, AUC) can be used. As with a study having the primary objective of defining a threshold or BMD departure point (section 3.1.3.4), it may be necessary to differentially allocate animals unequally to the various dose groups, e.g., to estimate the carcinogenic potential at human exposure levels it may be necessary use more than 50 animals in the low dose groups. A similar adjustment of animal numbers will have to be made in the chronic toxicity study. Determination of adverse effects (e.g., by use of the MTD) may be a lower priority in the study than maximizing low-dose sensitivity. Again, depending on the human exposure level, the middle portion of the dose–response curve may be reduced or excluded.

116. The data generated in such a study will be relevant to the prediction of effects at anticipated target population exposures. The power of the assay at the lower dose levels will be limited if the incidence of the responses of interest is very low. The approach is most relevant for frequently occurring health effects (e.g., respiratory irritation in occupational settings) where assay power is not particularly critical. However, such a study will have little or no value in identification of carcinogens.

117. It should also be noted that determination of the appropriate target population exposure level for use as the departure point for dose selection will be controversial; and, if actual (or future) exposure is determined to be higher than that used as point of departure in the bioassay design, the results will be of limited applicability. The target population exposure is likely to follow a distribution curve, perhaps a broad and highly skewed one. Thus, the adequacy of the study for the whole range of exposures may not be evident. Also, the results may not provide relevant information regarding the shape of the dose–response curve.

### 3.1.3.6 *Providing Data to Test Hypotheses Regarding Mode of Action of Carcinogens*

118. As noted in the Test Guidelines and as discussed in Chapter 2 of this Guidance Document, information on, and consideration of, the mode of action (MOA) of a suspected carcinogen ((Sonnich-Mullin et al., 2002; Cohen et al. 2003; Meek et al, 2003; Holsapple et al, 2006; Boobis et al., 2006, EPA, 2005) is particularly important, since the optimal design may differ depending on whether the substance is a known or suspected genotoxic carcinogen. These considerations may therefore dictate the design of a carcinogenicity study. In selecting dose levels for such a study, doses will need to be placed carefully, to yield observations of subtle precursor effects or other biomarkers of toxicity without inducing confounding effects related to frank toxicity. This approach requires some previously gathered information on potential modes of action, e.g. from genotoxicity studies.

119. The number of dose groups will be a critical element in the investigation and may need to be constrained in order to provide enough animals for assessment of multiple experimental points. Depending on the state of knowledge regarding possible modes of action (e.g., the number of possible alternatives), sensitivity may be a less important constraint on dose group number than the need for an experimental design that can compare the various alternative modes of interest. This is an approach driven by hypotheses regarding the mode of action.

120. Such a study design is of potentially great use for enhancing understanding of the way in which toxicity occurs and, by inference, what factors may influence the shape of the dose–response curve. The study may however provide less information about the shape of the dose–response curve in the observable range than those described above, and will be of reduced value for hazard identification. Consideration of dose selection placement is likely to be of limited value in obtaining information on mode of action, and more important considerations might be use of time-dependent dosing patterns, initiation/promotion protocols and auxiliary data on intermediate endpoints involved in the process of carcinogenesis, possibly obtained from separate experiments (Rhomberg et al., 2007).

### 3.1.3.7 *Integration of the objectives of a chronic bioassay*

121. Each different objective outlined above seeks to maximize the statistical power of the study at a different point on the dose–response curve. The focus may be on a level of response or on the shape and slope of the overall curve. The situation is also complicated by the fact that, below a certain dose, attempts to increase statistical power by redistributing animal numbers in particular dose groups become redundant. It should be noted also that a limitation to redistribution of animal numbers may complicate the interpretation of the findings and increase the cost of the study, since many more animals are required in order to increase the statistical power by a very small fraction.

122. The preceding sections have shown clearly that the design of a chronic bioassay is influenced by the primary objective of the study. Standardisation of the design across international boundaries or within different regulatory communities, however desirable from the point of view of mutual acceptance of data, for animal welfare and economic reasons, is in practice difficult to achieve, and the guidance provided on dose selection in sections 3.1.3.1 to 3.1.3.6 above is relevant in those cases where a particular objective has been specifically identified for the study.

123. In practice, however, only one bioassay is likely to be carried out, and the information to be obtained from it must satisfy a number of different objectives (Rhomberg et al., 2007). For the majority of bioassays, there will be one primary objective (typically that of identification of carcinogenic potential and/or chronic toxicity) and several subsidiary objectives such as characterizing the dose-response curve, extrapolating to low doses, or identifying a point of departure for a BMD. As stated by Rhomberg et al., 2007, the nature of the subsidiary objectives will be contingent on the intended outcome (Rhomberg et al., 2007). If a valid negative result is obtained in a carcinogenicity screening study, and this was the only objective of the study, there may be no further questions to be answered. If a positive result is obtained, however, a number of issues arise regarding the nature of the carcinogenic responses and their relevance to the levels of exposure of target populations, requiring further investigation into the nature and interpretation of the effects seen.

124. The study design selected at the outset should include dose levels that combine several objectives. As indicated by Rhomberg et al., one approach to achieve this is to include additional dose groups levels in a such a way that the optimal doses for a number of different objectives are all included (Rhomberg et al., 2007). Some doses would be optimized for some objectives and others for other objectives, essentially running several bioassays in tandem.

125. However, this is not feasible, given the animal welfare, economic and time constraints. When attempting to combine these various objectives into a single study, selection of dose levels must be done in a way that does not compromise the primary objective while still allowing a secondary objective to be pursued in an acceptable albeit suboptimal manner. There may be embellishments to the core design based on study objectives but it would be a rare event when an erosion of the core minimum would be acceptable. For example, a trade-off may be qualitative hazard identification and quantitative use of data to characterize dose response. The ILSI publications provide an in-depth discussion of the potential conflicts that arise from the differing objectives of a bioassay (ILSI, 1997, Rhomberg et al., 2007).

126. As indicated by Rhomberg et al, 2007, a compromise between case-by-case optimization and a standard design for dose selection is to focus on four core selection schemes, as presented below:

1. **Hazard Screening Plus Dose–Response:** This is modeled on the current carcinogenicity bioassay. The top dose is chosen to increase the study’s statistical power to detect effects that may be rare. A second dose combines two functions: (1) hedging against the top dose being found to have been too high in retrospect, and (2) providing the opportunity for dose–response characterization of any effects found. Other lower dose levels can be placed so as to inform dose–response, no-effect levels, or other purposes. Key challenges will be balancing statistical power and toxicological relevance of the high dose level and compromising among subsidiary objectives while accounting for relevant dose-related physiological changes when setting lower dose levels.
2. **NOAEL/BMD-Seeking for Threshold Effects:** This is modeled on the current chronic study for non-cancer effects. The main aim is to identify no-effect (or low-effect) levels for the more sensitive adverse threshold effects. The top dose should aim at engendering adverse effects, the lowest dose should aim at constituting a NOAEL, and intermediate doses should be set so as to identify the dose levels at which the high-dose responses become manifest.

3. **Assessment of Safety of Human Exposure Levels:** This is modeled on safety assessment studies for nutrients and pharmaceuticals. For agents that are not genotoxic, show low toxicity, and evince no known difference in metabolic profile between rodents and humans, one can test multiples of anticipated human exposure. Lack of adverse effects at doses sufficiently above human exposure (and the perceived implausibility of non-threshold effects) gives evidence supporting the safety of the anticipated exposures. The bioassay exposures should be selected on an appropriate basis for animal:human comparison; for instance, the application to pharmaceuticals is typically based on area under the blood concentration-time curve that results from anticipated human exposures.
4. **Special-Purpose Bioassays:** Whenever the main emphases of the above three core schemes do not apply—or when they are dominated by another compelling purpose—it is necessary to consider a more completely case-specific design. For instance, if one is conducting a second bioassay to address dose–response properties and organ toxicity dependence of a tumor response that has previously been discovered in a screening bioassay, the dose selection should be optimized for the specific purposes at hand, and the choice of the top dose level will no longer be dominated by hazard identification concerns. In such situations, the study design is not likely to follow any of the designs outlined above for chronic toxicity and/or carcinogenicity studies, but will be designed to answer a specific question as already indicated.

## REFERENCES

Boobis, A.R., Cohen, S.M., Dellarco, V., McGregor, D., Meek, M.E., Vickers, C., Willcocks, D. & Farland, W. IPCS Framework for analyzing the Relevance of a Cancer Mode of Action for Humans. *Crit. Rev. in Toxicol.* (2006) 36:793-801.

Cohen, S.M., Meek, M.E., Klaunig, J.E., Patton, D.E., and Fenner-Crisp, P.A. (2003) The human relevance of information on carcinogenic Modes of Action: An Overview. *Crit. Rev. Toxicol.* 33:581-589

ECETOC. 1996. Practical concepts for dose selection in chronic toxicity and carcinogenicity studies in rodents. ECETOC Monograph No. 25. European Centre for Ecotoxicology and Toxicology of Chemicals. Brussels, Belgium.

EPA (2005). Guidelines for Carcinogen Risk Assessment Risk Assessment Forum U.S. Environmental Protection Agency Washington, DC

Holsapple, M.P., Pitot, H.C., Cohen, S.N., Boobis, A.R., Klaunig, J.E., Pastoor, T., Dellarco, V.L. & Dragan, Y.P. (2006) Mode of Action in Relevance of Rodent Liver Tumors to Human Cancer Risk. *Toxicol. Sci.* 89:51-56.

ICH (2008). Guidance for Industry . S1C(R2) Dose Selection for Carcinogenicity Studies U.S. Department of Health and Human Services Food and Drug Administration Center for Drug Evaluation and Research (CDER) Center for Biologics Evaluation and Research (CBER). September 2008 ICH. Revision 1

ILSI (International Life Sciences Institute) (1997). Principles for the Selection of Doses in Chronic Rodent Bioassays. Foran JA (Ed.). ILSI Press, Washington, DC.

Meek, E.M., Bucher, J.R., Cohen, S.M., Dellarco, V., Hill, R.N., Lehman-McKemmon, L.D., Longfellow, D.G., Pastoor, T., Seed, J. & Patton, D.E. A Framework for Human Relevance analysis of Information on Carcinogenic Modes of Action. *Crit. Rev. Toxicol.* (2003). 33:591-653.

Rhomberg, L., Baetcke, K., Blancato, J., Bus, J., Cohen, S., Conolly, R., Dixit, R., Doe, J., Ekelman, K., Fenner-Crisp, P., Harvey, P., Hattis, D., Jacobs, A., Jacobson-Kram, D., Lewandowski, T., Liteplo, R., Pelkonen, O., Rice, J., Somers, D, Turturro, A., West, W. Olin, S. (2007). Issues in the Design and Interpretation of Chronic Toxicity and Carcinogenicity Studies in Rodents: Approaches to Dose Selection. *Crit. Rev. Toxicol.*37:729-837.

Smith, G. (2002). Scientific review of chronic rodent bioassays. Criteria for Maximum Tolerated Dose (MTD) attainment: clinical pathology- and pathology-related endpoints (Unpublished draft, submitted to Pest Management Regulatory Agency, Canada, September 18, 2002).

Sonnich-Mullin, C., Fielder, R., Wiltse, J., Baetcke, K., Dempsey, J., Fenner-Crisp, P., Grant, D., Hartley, M., Knaap, A., Kroese, D., Mangelsdorf, I., Meek, E., Rice, J., & Younes, M. IPCS Conceptual Framework for Evaluating a Mode of Action for Chemical Carcinogenesis. *Reg. Toxicol. & Pharm.* (2001) 34:146-152.

## **Appendix 1: ILSI Principles for Dose Selection in Chronic Rodent Bioassays**

### *Principle 1*

Dose selection for chronic studies must be based on sound toxicologic principles. Within a reasonable dose range, increasing the dose can increase the ability to detect an effect; therefore, doses for chronic rodent bioassays should be selected within this range to maximize the sensitivity of a chronic bioassay. However, trying to increase study sensitivity by increasing doses into ranges that do not reflect application of sound toxicologic principles could lead to results that are inappropriate for human risk assessment.

Increasing the highest dose in a chronic bioassay may increase sensitivity within some defined dose range, but the potential exists that different mechanisms of toxicity or chemical mode of action are active at higher doses, which may not be relevant to humans exposed to lower doses. In this case, selection of the highest dose may be influenced by consideration of the mechanism/ mode of action and other factors discussed in Principle 4. However, when the highest dose in a carcinogenicity assay is limited by effects (e.g., a mode of action in one organ system) that are thought not to occur in humans, one must be aware that it still is possible that a higher dose of the chemical may be carcinogenic in other animal/organ systems.

To address these issues, the working group encourages an approach to dose selection that incorporates all relevant information from prechronic studies and other sources, uses toxicologic tools associated with an understanding of the mechanisms or mode of action by which a chemical produces an effect (e.g., genotoxicity, cell proliferation, etc.), and uses good scientific principles to enhance the accuracy of judgments of potential human risks. In the case of negative studies (particularly where the highest dose is chosen based on a full characterization of the chemical's toxicity in prechronic studies), use of sound scientific principles as well as all available chemical, physical, and toxicologic data will lessen concern that the result may be a false negative. Similarly, in positive carcinogenicity studies, this approach will lessen concern that the result may be a false positive. In both cases, the predictiveness of the bioassay for human health effects will be improved.

### *Principle 2*

A chronic bioassay requires a major investment in resources and time, and the objective of such a study should be broader than hazard identification. Scientists who conduct chronic bioassays and those who use data from bioassays, including regulatory agencies, should encourage innovative approaches to dose selection by considering appropriate study designs, mechanistic data, and other information in the design and interpretation of studies. Use of additional endpoints and other information must be based on sound scientific rationale, and such designs should be evaluated on the basis of their individual merits.

A goal of high-dose selection in carcinogenicity bioassays is, in the context of hazard identification, to reduce the likelihood of a false-negative result. However, it is recognized that the qualitative nature of the hazard (e.g., carcinogenic response) may itself be dose dependent. This principle encourages approaches to dose selection that incorporate consideration of mechanistic and other toxicologic information. Such approaches should improve the scientific basis for dose selection and aid in interpretation of data generated from chronic bioassays.

### *Principle 3*

Human exposure should be considered in dose selection, particularly for selection of the middle and lowest doses. Further, the middle and lowest doses should be selected to characterize the shape of the dose response curve as much as possible. Selection of the middle and lower doses should take into account

factors such as the mechanism or mode of action, toxicokinetics, and others listed in Principles 4 and 5 and should not be based solely on a fraction of the highest dose.

Issues that should be considered when incorporating potential human exposure in dose selection include the human exposure route and mode, the dose range in the chronic bioassay in relation to human exposure, and the duration and frequency of human exposure, if known. Subpopulations that may be more highly exposed than the general population, or that are genetically more susceptible, also should be considered. The relationship between external and delivered (internal) dose (e.g., ingested dose versus dose delivered to the target organ, toxicokinetics) in both humans and test organisms may influence dose selection. Further, for substances expected to exhibit a toxicity threshold, or if the evaluation of carcinogenic potential is being combined with an evaluation of chronic toxicity, the study should be designed to include one dose that does not elicit adverse effects; that is, one dose should be a NOAEL. Of course, caution must be exercised to ensure that the NOAEL is not simply an artifact of small sample size or poor study design.

#### *Principle 4*

The [ILSI] working group has recommended the use of innovative approaches, additional endpoints, and other information in the selection of doses for chronic rodent bioassays. The following endpoints, generally determined in prechronic studies, should be considered in dose selection for chronic rodent bioassays. Further, it is recognized that endpoints other than those listed below may provide important information for dose selection, and use of those endpoints, where they are based on sound toxicologic principles, is encouraged. Such endpoints may be available presently, or they may be developed as the science of toxicology advances.