Unclassified ENV/JM/MONO(2007)28

Organisation de Coopération et de Développement Economiques
Organisation for Economic Co-operation and Development

26-Sep-2007

English - Or. English

ENVIRONMENT DIRECTORATE
JOINT MEETING OF THE CHEMICALS COMMITTEE AND
THE WORKING PARTY ON CHEMICALS, PESTICIDES AND BIOTECHNOLOGY

ENV/JM/MONO(2007)28
Unclassified

SERIES ON TESTING AND ASSESSMENT
Number 80

GUIDANCE ON GROUPING OF CHEMICALS

JT03232745

English - Or. English

OECD Environment Health and Safety Publications

Series on Testing and Assessment

No. 80

GUIDANCE ON GROUPING OF CHEMICALS

IOMC

INTER-ORGANIZATION PROGRAMME FOR THE SOUND MANAGEMENT OF CHEMICALS

A cooperative agreement among UNEP, ILO, FAO, WHO, UNIDO, UNITAR and OECD

Environment Directorate

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

Paris 2007

Also published in the Series on Testing and Assessment:

No. 1, *Guidance Document for the Development of OECD Guidelines for Testing of Chemicals (1993; reformatted 1995, revised 2006)*

No. 2, *Detailed Review Paper on Biodegradability Testing (1995)*

No. 3, *Guidance Document for Aquatic Effects Assessment (1995)*

No. 4, *Report of the OECD Workshop on Environmental Hazard/Risk Assessment (1995)*

No. 5, *Report of the SETAC/OECD Workshop on Avian Toxicity Testing (1996)*

No. 6, *Report of the Final Ring-test of the Daphnia magna Reproduction Test (1997)*

No. 7, *Guidance Document on Direct Phototransformation of Chemicals in Water (1997)*

No. 8, *Report of the OECD Workshop on Sharing Information about New Industrial Chemicals Assessment (1997)*

No. 9, *Guidance Document for the Conduct of Studies of Occupational Exposure to Pesticides during Agricultural Application (1997)*

No. 10, *Report of the OECD Workshop on Statistical Analysis of Aquatic Toxicity Data (1998)*

No. 11, *Detailed Review Paper on Aquatic Testing Methods for Pesticides and industrial Chemicals (1998)*

No. 12, *Detailed Review Document on Classification Systems for Germ Cell Mutagenicity in OECD Member Countries (1998)*

No. 13, *Detailed Review Document on Classification Systems for Sensitising Substances in OECD Member Countries 1998)*

No. 14, *Detailed Review Document on Classification Systems for Eye Irritation/Corrosion in OECD Member Countries (1998)*

No. 15, *Detailed Review Document on Classification Systems for Reproductive Toxicity in OECD Member Countries (1998)*

No. 16, *Detailed Review Document on Classification Systems for Skin Irritation/Corrosion in OECD Member Countries (1998)*

No. 17, *Environmental Exposure Assessment Strategies for Existing Industrial Chemicals in OECD Member Countries (1999)*

No. 18, *Report of the OECD Workshop on Improving the Use of Monitoring Data in the Exposure Assessment of Industrial Chemicals (2000)*

No. 19, *Guidance Document on the Recognition, Assessment and Use of Clinical Signs as Humane Endpoints for Experimental Animals used in Safety Evaluation (1999)*

No. 20, *Revised Draft Guidance Document for Neurotoxicity Testing (2004)*

No. 21, *Detailed Review Paper: Appraisal of Test Methods for Sex Hormone Disrupting Chemicals (2000)*

No. 22, *Guidance Document for the Performance of Out-door Monolith Lysimeter Studies (2000)*

No. 23, *Guidance Document on Aquatic Toxicity Testing of Difficult Substances and Mixtures (2000)*

No. 24, *Guidance Document on Acute Oral Toxicity Testing (2001)*

No. 25, *Detailed Review Document on Hazard Classification Systems for Specifics Target Organ Systemic Toxicity Repeated Exposure in OECD Member Countries (2001)*

No. 26, *Revised Analysis of Responses Received from Member Countries to the Questionnaire on Regulatory Acute Toxicity Data Needs (2001)*

No 27, *Guidance Document on the Use of the Harmonised System for the Classification of Chemicals Which are Hazardous for the Aquatic Environment (2001)*

No 28, *Guidance Document for the Conduct of Skin Absorption Studies (2004)*

No 29, *Guidance Document on Transformation/Dissolution of Metals and Metal Compounds in Aqueous Media (2001)*

No 30, *Detailed Review Document on Hazard Classification Systems for Mixtures (2001)*

No 31, *Detailed Review Paper on Non-Genotoxic Carcinogens Detection: The Performance of In-Vitro Cell Transformation Assays (2007)*

No. 32, *Guidance Notes for Analysis and Evaluation of Repeat-Dose Toxicity Studies (2000)*

No. 33, *Harmonised Integrated Classification System for Human Health and Environmental Hazards of Chemical Substances and Mixtures (2001)*

No. 34, *Guidance Document on the Development, Validation and Regulatory Acceptance of New and Updated Internationally Acceptable Test Methods in Hazard Assessment (2005)*

No. 35, *Guidance notes for analysis and evaluation of chronic toxicity and carcinogenicity studies (2002)*

No. 36, *Report of the OECD/UNEP Workshop on the use of Multimedia Models for estimating overall Environmental Persistence and long range Transport in the context of PBTS/POPS Assessment (2002)*

No. 37, *Detailed Review Document on Classification Systems for Substances Which Pose an Aspiration Hazard (2002)*

No. 38, *Detailed Background Review of the Uterotrophic Assay Summary of the Available Literature in Support of the Project of the OECD Task Force on Endocrine Disrupters Testing and Assessment (EDTA) to Standardise and Validate the Uterotrophic Assay (2003)*

No. 39, *Guidance Document on Acute Inhalation Toxicity Testing (in preparation)*

No. 40, *Detailed Review Document on Classification in OECD Member Countries of Substances and Mixtures Which Cause Respiratory Tract Irritation and Corrosion (2003)*

No. 41, *Detailed Review Document on Classification in OECD Member Countries of Substances and Mixtures which in Contact with Water Release Toxic Gases (2003)*

No. 42, *Guidance Document on Reporting Summary Information on Environmental, Occupational and Consumer Exposure (2003)*

No. 43, *Draft Guidance Document on Reproductive Toxicity Testing and Assessment (in preparation)*

No. 44, *Description of Selected Key Generic Terms Used in Chemical Hazard/Risk Assessment (2003)*

No. 45, *Guidance Document on the Use of Multimedia Models for Estimating Overall Environmental Persistence and Long-range Transport (2004)*

No. 46, *Detailed Review Paper on Amphibian Metamorphosis Assay for the Detection of Thyroid Active Substances (2004)*

No. 47, *Detailed Review Paper on Fish Screening Assays for the Detection of Endocrine Active Substances (2004)*

No. 48, *New Chemical Assessment Comparisons and Implications for Work Sharing (2004)*

No. 49, *Report from the Expert Group on (Quantitative) Structure-Activity Relationships [(Q)SARs] on the Principles for the Validation of (Q)SARs (2004)*

No. 50, *Report of the OECD/IPCS Workshop on Toxicogenomics (2005)*

No. 51, *Approaches to Exposure Assessment in OECD Member Countries: Report from the Policy Dialogue on Exposure Assessment in June 2005 (2006)*

No. 52, *Comparison of emission estimation methods used in Pollutant Release and Transfer Registers (PRTRs) and Emission Scenario Documents (ESDs): Case study of pulp and paper and textile sectors (2006)*

No. 53, *Guidance Document on Simulated Freshwater Lentic Field Tests (Outdoor Microcosms and Mesocosms) (2006)*

No. 54, *Current Approaches in the Statistical Analysis of Ecotoxicity Data: A Guidance to Application (2006)*

No. 55, *Detailed Review Paper on Aquatic Arthropods in Life Cycle Toxicity Tests with an Emphasis on Developmental, Reproductive and Endocrine Disruptive Effects (2006)*

No. 56, *Guidance Document on the Breakdown of Organic Matter in Litter Bags (2006)*

No. 57, *Detailed Review Paper on Thyroid Hormone Disruption Assays (2006)*

No. 58, *Report on the Regulatory Uses and Applications in OECD Member Countries of (Quantitative) Structure-Activity Relationship [(Q)SAR] Models in the Assessment of New and Existing Chemicals (2006)*

No. 59, *Report of the Validation of the Updated Test Guideline 407: Repeat Dose 28-Day Oral Toxicity Study in Laboratory Rats (2006)*

No. 60, *Report of the Initial Work Towards the Validation of the 21-Day Fish Screening Assay for the Detection of Endocrine Active Substances (Phase 1A) (2006)*

No. 61, *Report of the Validation of the 21-Day Fish Screening Assay fort he Detection of Endocrine Active Substances (Phase 1B) (2006)*

No. 62, *Final OECD Report of the Initial Work Towards the Validation of the Rat Hershberger Assay : Phase-1, Androgenic Response to Testosterone Propionate, and Anti-Androgenic Effects of Flutamide (2006)*

No. 63, *Guidance Document on the Definition of Residue (2006)*

No. 64, *Guidance Document on Overview of Residue Chemistry Studies (2006)*

No. 65, *OECD Report of the Initial Work Towards the Validation of the Rodent Utertrophic Assay - Phase 1 (2006)*

No. 66, *OECD Report of the Validation of the Rodent Uterotrophic Bioassay: Phase 2. Testing of Potent and Weak Oestrogen Agonists by Multiple Laboratories (2006)*

No. 67, *Additional data supporting the Test Guideline on the Uterotrophic Bioassay in rodents (draft)*

No. 68, *Summary Report of the Uterotrophic Bioassay Peer Review Panel, including Agreement of the Working Group of the National Coordinators of the Test Guidelines Programme on the follow up of this report (2006)*

No. 69, *Guidance Document on the Validation of (Quantitative) Structure-Activity Relationship [(Q)SAR] Models (2007)*

No. 70, *Report on the Preparation of GHS Implementation by the OECD Countries (2007)*

No. 71, *Guidance Document on the Uterotrophic Bioassay - Procedure to Test for Antioestrogenicity (2007)*

No. 72, *Guidance Document on Pesticide Residue Analytical Methods (2007)*

No. 73, *Report of the Validation of the Rat Hershberger Assay: Phase 3: Coded Testing of Androgen Agonists, Androgen Antagonists and Negative Reference Chemicals by Multiple Laboratories. Surgical Castrate Model Protocol (2007)*

No. 74, *Detailed Review Paper for Avian Two-generation Toxicity Testing (2007)*

No. 75, *Guidance Document on the Honey Bee (Apis Mellifera L.) Brood test Under Semi-field Conditions (2007)*

No. 76, *Final Report of the Validation of the Amphibian Metamorphosis Assay for the Detection of Thyroid Active Substances: Phase 1 - Optimisation of the Test Protocol (2007)*

No. 77, *Final Report of the Validation of the Amphibian Metamorphosis Assay: Phase 2 - Multi-chemical Interlaboratory Study (2007)*

No. 78, *Final report of the Validation of the 21-day Fish Screening Assay for the Detection of Endocrine Active Substances. Phase 2: Testing Negative Substances (2007)*

No. 79, *Validation Report of the Full Life-cycle Test with the Harpacticoid Copepods Nitocra Spinipes and Amphiascus Tenuiremis and the Calanoid Copepod Acartia Tonsa - Phase 1 (2007)*

No. 80, *Guidance on Grouping of Chemicals (2007)*

About the OECD

The Organisation for Economic Co-operation and Development (OECD) is an intergovernmental organisation in which representatives of 30 industrialised countries in North America, Europe and the Asia and Pacific region, as well as the European Commission, meet to co-ordinate and harmonise policies, discuss issues of mutual concern, and work together to respond to international problems. Most of the OECD's work is carried out by more than 200 specialised committees and working groups composed of member country delegates. Observers from several countries with special status at the OECD, and from interested international organisations, attend many of the OECD's workshops and other meetings. Committees and working groups are served by the OECD Secretariat, located in Paris, France, which is organised into directorates and divisions.

The Environment, Health and Safety Division publishes free-of-charge documents in ten different series: **Testing and Assessment**; **Good Laboratory Practice and Compliance Monitoring**; **Pesticides and Biocides**; **Risk Management**; **Harmonisation of Regulatory Oversight in Biotechnology**; **Safety of Novel Foods and Feeds; Chemical Accidents; Pollutant Release and Transfer Registers; Emission Scenario Documents; and the Safety of Manufactured Nanomaterials.** More information about the Environment, Health and Safety Programme and EHS publications is available on the OECD's World Wide Web site (http://www.oecd.org/ehs/).

This publication was produced within the framework of the Inter-Organisation Programme for the Sound Management of Chemicals (IOMC).

The Inter-Organisation Programme for the Sound Management of Chemicals (IOMC) was established in 1995 following recommendations made by the 1992 UN Conference on Environment and Development to strengthen co-operation and increase international co-ordination in the field of chemical safety. The participating organisations are FAO, ILO, OECD, UNEP, UNIDO, UNITAR and WHO. The World Bank and UNDP are observers. The purpose of the IOMC is to promote co-ordination of the policies and activities pursued by the Participating Organisations, jointly or separately, to achieve the sound management of chemicals in relation to human health and the environment.

This publication is available electronically, at no charge.

For this and many other Environment, Health and Safety publications, consult the OECD's World Wide Web site (www.oecd.org/ehs/)

or contact:

OECD Environment Directorate,
Environment, Health and Safety Division

2 rue André-Pascal
75775 Paris Cedex 16
France

Fax: (33-1) 44 30 61 80

E-mail: ehscont@oecd.org

FOREWORD

This guidance document is part of the OECD effort to provide guidance for assessing the hazards of chemical substances while gaining efficiencies and improving animal welfare. The approach described in this guidance document is to consider closely related chemicals as a group, or category, rather than as individual chemicals. In the category approach, not every chemical needs to be tested for every endpoint. Rather, the overall data for that category must prove adequate to support a hazard assessment. The overall data set must allow the estimation of the hazard for the untested endpoints.

Although this approach has been used on an *ad hoc* basis in many regulatory programmes for many years, a guidance document was first developed by the US-EPA in support of the US HPV Challenge Program in 1998. The same guidance document was also inserted into the *OECD Manual for Investigation of HPV Chemicals*. Since then the guidance has evolved continuously based on experience with the approach within the OECD HPV Chemicals Programme as well as national/regional regulatory and voluntary frameworks. The publication of this guidance document in the Series on Testing and Assessment of the OECD Environment, Health and Safety Publications is aimed at improving the visibility of this approach and at recommending its wider use. It is nevertheless recognised that the technique of assessing groups of substances is an evolving field and that continuous revisions of this guidance document is envisaged. Furthermore, due to the evolving nature of the approach as well as its complexity, early consultations between industry and authorities is recommended when using it for regulatory purposes.

The guidance first explains what a category is and relevant concepts that will enable the document to be better read (Chapter 2). In this chapter the mechanistic basis for categories is explained and the advantages derived from using a category described. Chapter 2 also describes the close relationship that exists between (Q)SAR and categories, both in terms of the concepts and in the use of (Q)SAR for data evaluation and gap-filling. Chapter 3 describes the main approaches that are used for data gap filling: read-across, trend analysis and QSARs. While Chapters 2 and 3 provide explanations on the scientific and methodological background of the analogue and category approaches, respectively, Chapters 4-7 focus more on practical aspects for forming and documenting analogue and chemical category approaches. Separate chapters (4 and 5) were elaborated to provide guidance on stepwise procedures for analogue read-across and chemical categories, so that the guidance document can be used in a “modular” fashion, making it possible to use parts of the guidance only. Therefore a number of repetitions of texts were also necessary. Chapter 6 elaborates on some specific issues that need to be addressed with specific types of categories. Finally, in Chapter 7, a Category Reporting Format is proposed as a tool for documenting chemical categories.

TABLE OF CONTENTS

CHAPTER 1. INTRODUCTION

There are many national, regional and international programmes – either regulatory or voluntary – to assess the hazards or risks of chemical substances to humans and the environment. The first step in making a hazard assessment of a substance is to ensure that there is adequate information on each of the endpoints. If adequate information is not available then additional data is needed to complete the dataset for this substance.

For reasons of resources and animal welfare, it is important to limit the number of tests to be conducted, where this is scientifically justifiable. One approach is to consider closely related chemicals as a group, or chemical category, rather than as individual chemicals. In the category approach, not every chemical needs to be tested for every required endpoint. Rather, the data for chemicals and endpoints that have been tested are used to estimate the corresponding properties for the untested chemicals and endpoints. The overall category data and rationale need be adequate to support a screening-level hazard assessment.

Another approach to limiting animal testing is to use an analogue approach[1] where comparisons are made between a very limited number of chemicals. Endpoint information for one chemical is used to predict the same endpoint for another chemical, which is considered to be "similar" in some way (usually on the basis of structural similarity and similar properties and/or activities). This simple approach is generally open to more uncertainty than the broader category approach.

An additional advantage of a chemical category assessment approach is that identification of consistent patterns of effects within a category in itself increases confidence in the reliability of the results for all the individual substances in the category, compared to evaluation of data purely on a substance-by-substance basis.

All assessments should be reviewed and updated as new information is generated, because category assessments are often complex and experience in forming and assessing categories is growing. Therefore, periodic review and update of category assessments provides a means of incorporating new information, re-affirming or strengthening the scientific basis of the original premise for the category, and ensuring that the methodology associated with category assessments is continually improved. There may be cases where new information is generated for a category member which calls the category justification into question. In such cases, the category should be re-evaluated and may need to be re-constructed (further guidance is available in Chapter 5)

This document has been developed based on existing cases involving chemical categories assessed within the OECD HPV Chemicals Programme, the US HPV Challenge Programme, the EU Existing Substances Programme, the EU activity on classification and labelling, guidance issued under the

[1] In this document, the term "analogue approach" is used to describe the assessment of small numbers of similar chemicals and "read-across" to describe a way of filling data gaps.

US HPV Challenge Programme and other US EPA programmes as well as for the EU REACH legislation, and on the experience gained from the OECD Workshop on the development and use of chemical categories held in January 2004. Furthermore, this document addresses the actual formation of categories for test plan and hazard assessment purposes, and it makes some preliminary suggestions about presentation. The document will need to be updated as further experience is gained.

The regulatory application of QSAR methods for providing data for specific endpoints is outside of the scope of this document and can be found in the following documents:

- Section 3.3 of the OECD Manual for Investigation of HPV Chemicals provides guidance on the use of SAR in the HPV Chemicals Programme (OECD, 2007b)

- an OECD Report on the Regulatory Uses and Applications in OECD Member Countries of (Q)SAR Models in the Assessment of New and Existing Chemicals (OECD, 2006a) summarises the experience of OECD member countries with QSAR applications

- an OECD report on the principles for the validation, for regulatory purposes, of (Q)SAR models (OECD, 2004) and an accompanying OECD guidance document (OECD, 2007a).

CHAPTER 2. EXPLANATION OF THE CHEMICAL CATEGORY APPROACH

2.1 Introduction

In this guidance document, the terms category approach and analogue approach are used to describe techniques for grouping chemicals, whilst the term read-across is reserved for a technique of filling data gaps in either approach. A chemical category is a group of chemicals whose physico-chemical and human health and/or environmental toxicological properties and/or environmental fate properties are likely to be similar or follow a regular pattern as a result of structural similarity (or other similarity characteristic). In principle, more members are generally present in a chemical category, enabling the detection of trends across endpoints. As the number of possible chemicals being grouped into a category increases, the potential for developing hypotheses for specific endpoints and making generalisations about the trends within the category will also increase, and hence increase the robustness of the evaluation. The term analogue approach is used when the grouping is based on a very limited number of chemicals, where trends in properties are not apparent.Categories of chemicals are selected based on the hypothesis that the properties of a series of chemicals with common structural features will show coherent trends in their physicochemical properties, and more importantly, in their toxicological (human health/ecotoxicity) effects or environmental fate properties. Common behaviour or consistent trends are generally associated with a common underlying mechanism of action, or where a mechanism of action exhibits intensity changes in a consistent manner across the different members of a category.

The use of a category approach will mean that it is possible to identify properties which are common to at least some members of the category. The approach also provides a basis on which to identify possible trends in properties across the category. As a result, it is possible to extend the use of measured data to similar untested chemicals, and reliable estimates that are adequate for classification and labelling and/or risk assessment can be made without further testing. In addition, knowledge of the expected effects of the category together with information on use and exposure will help in deciding not only whether additional testing is needed, but also the nature and scope of any testing that needs to be carried out.

The assessment of chemicals by using a category approach differs from the approach of assessing them on an individual basis, since the effects of the individual chemicals within a category are assessed on the basis of the evaluation of the category as a whole, rather than based on measured data for any one particular substance alone. For a category member that lacks data for an endpoint, the data gap can be filled in a number of ways, including by read-across from one or more other category members. In some circumstances, it may only be necessary to use data from one category member using read-across principles to adequately characterize the member lacking data. The category approach is important since it provides an alternative to testing individual substances and as a result should lead to a decrease in the use of animal testing.

2.2 Benefits of the chemical category approach

Assessment of a large number of chemicals as a category can be more efficient and accurate than assessment of single compounds for a number of reasons:

a) data from one or more chemicals can be interpolated or extrapolated to other chemicals, reducing the need to test for every endpoint for every chemical;
b) since existing data can be applied to additional chemicals without the need for additional testing, the use of animal testing is reduced;
c) the category evaluation is based on a greater body of data than on data on a single compound;
d) the identification of compounds as members of a category provides an insight into the potential effects of the compounds that might otherwise be overlooked;
e) the use of a category approach may also provide significant advantages in the evaluation of compounds that are often considered as "difficult", in the sense that they can present technical difficulties when carrying out standard test protocols (examples are given in Hart J, 2007; Comber M & Simpson B, 2007);
f) the approach provides a valuable tool in cases where animal models do not always reliably predict effects on humans (examples are given in Hart J, 2007),
g) in most cases, category testing can be completed earlier than individual tests for each chemical that requires notification, submission or inclusion,
h) in order to gain future efficiencies, category proposals may be expanded via the inclusions of chemicals that may be addressed under various global programs,
i) in the category approach, not every chemical needs to be tested for every endpoint. Rather, the overall data for that category must prove adequate to support a hazard assessment. The overall data set must allow the estimation of the hazard for the missing data points,
j) a category approach allows for better consideration of the biological plausibility of grouping the chemicals within a category.

Use of a category approach can also provide significant efficiencies and benefits when identifying data gaps and filling data needs that are ultimately deemed necessary. A category test plan is designed to provide information to characterise the group as a whole rather than to fill every data point for every chemical in the category. This reflects an approach that is more efficient from a testing perspective than test plans for obtaining data on individual chemicals of commercial interest. A knowledge of the expected biological effects of the category will be helpful in deciding not only whether testing is needed, but also the nature and scope of the test to be carried out. Where confirmation is sought that an individual category member does not have a particular property (e.g. acute oral toxicity), a simple limit test might be adequate to provide the necessary confirmation. Where an individual category member is expected to have an effect (e.g. skin irritation or corrosion), a simple *in vitro* test might provide adequate confirmation of the predicted effect.

Another benefit of using a category approach is that this approach allows for an evaluation of the biological basis for the effects seen in a group of chemicals within a category. When it is known that members of a chemical category share a presumed common mechanism of action, the confidence in the category is significantly greater than that associated with the use of a read-across approach to fill data gaps. This confidence increases with increasing numbers of chemicals included in the category. For a large category[2], both the presence and absence of certain hazards, as well as the trend of an effect across a category, can be identified. This provides a basis on which the properties of individual members of the

[2] Based on the current experience within the OECD HPV Chemicals Programme, any category with more than 10 members is a large category.

category can be identified with the necessary confidence. For more limited comparisons, particularly with chemicals containing multiple functional groups, it is harder to obtain the same level of confidence. A category approach can provide significant advantages compared to the read-across techniques for filling data gaps, in that it is possible to analyse trends in properties. Read-across techniques between chemical analogues have been extensively used (e.g. within the OECD HPV Chemicals Programme, the EU Existing Chemicals Programme or for Classification and Labelling in the EU), often on an *ad hoc* basis and it is foreseen that they will continue to be used extensively. Nevertheless, an important consideration in preparing this Guidance is to encourage the replacement of these *ad hoc* approaches by a more wide-ranging approach that can provide a greater degree of confidence in the result.

Guidance on the analogue approach is provided in Chapter 4, and guidance on category formation is provided in Chapter 5.

2.3 Explanation of relevant concepts

The term 'grouping' or 'chemical grouping' describes the general approach to assessing more than one chemical at the same time. It can include formation of a chemical category or identification of a chemical analogue for which read-across may be applied. In this document, the more specific terms 'chemical category' and 'analogue approach' are used.

A chemical category is a group of chemicals whose physicochemical and human health and/or environmental toxicological properties and/or environmental fate properties are likely to be similar or follow a regular pattern as a result of structural similarity. The similarities may be based on the following:

a) common functional group(s) (e.g. aldehyde, epoxide, ester, specific metal ion);
b) common constituents or chemical classes, similar carbon range numbers. This is frequently the case with complex substances often known as "substances of Unknown or Variable composition, Complex reaction products or Biological material" (UVCB substances);
c) an incremental and constant change across the category (e.g. a chain-length category), often observed in physicochemical properties, e.g. boiling point range;
d) the likelihood of common precursors and/or breakdown products, via physical or biological processes, which result in structurally similar chemicals (e.g. the "metabolic pathway approach" of examining related chemicals such as acid/ester/salt).

Categories can be developed systematically on the basis of structure (or other similar characteristic) alone. It is recognised that in many cases the formation of a chemical category is also dependant on which chemicals are manufactured by the consortium of companies sponsoring the category and/or the regulatory context under which the evaluation is being made. While these considerations can legitimately influence the formation of a category, they are independent of the scientific analysis of a category.

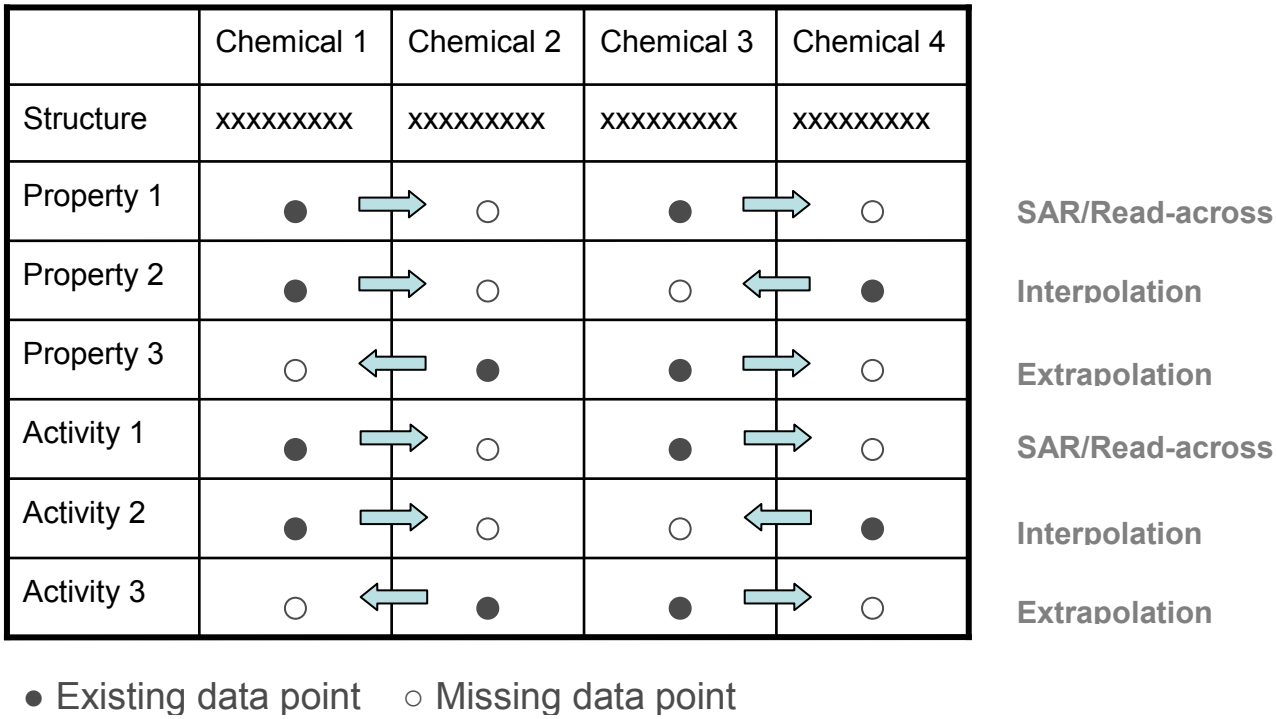
Within a chemical category, data gaps may be filled by read-across, trend analysis and QSARs. Read-across is a technique used to predict endpoint information for one chemical by using data from the same endpoint from another chemical which is considered to be 'similar' in some way (on the basis of structural similarity and similar properties and/or activities). For a given category endpoint, the category members are often related by a trend (e.g. increasing, decreasing or constant) in an effect, and a trend analysis can be carried out using a model based on the data for the members of the category. Data gaps can also be filled by an external QSAR model, where the category under examination is a subcategory of the wider QSAR. Further details are given in Chapter 3.

While read-across is a technique for data gap filling within the context of a category approach, it is also a useful tool for data gap filling in cases where comparisons are based on a very limited number of chemicals. The simplest example of the category approach is a comparison between two chemicals. This form of evaluation is often called a read-across approach. This approach has been used extensively in the evaluation of chemicals under a number of different evaluation programmes, and, although the approach has been used on a largely *ad hoc* basis, there are a number of examples on which guidance can be based. Whilst sharing many characteristics in common with a category approach, the evaluation of a very limited number of chemicals does present a number of differences compared to the evaluation of larger, systematically derived categories, for which there is more limited experience. In order to avoid confusion, evaluations of a very limited number of chemicals using largely read-across to fill data gaps is described in this guidance as the analogue approach. The term read-across is therefore limited to the technique for filling data gaps described in Chapter 3.

In the analogue approach, endpoint information for one chemical is used to predict the same endpoint for another chemical, which is considered to be "similar" in some way (usually on the basis of structural similarity and similar properties and/or activities). General guidance on how to use the analogue approach is provided in Chapter 4.

A chemical category can be described by a matrix consisting of the category members and by a corresponding set of properties and/or effects data (the category endpoints), (see Figure 1). General guidance on how to build categories is provided in Chapter 5, whereas specific guidance for different types of categories is given in Chapter 6.

Figure 1. Graphical representation of a chemical category and some approaches for filling data gaps



As illustrated in Figure 1, data gap filling can be done using read-across from one tested chemical to an untested chemical. In general, interpolation is preferred to extrapolation between category members, this is discussed in more detail in Chapter 3, section 3.3. Other approaches which include trend analysis, (Q)SARs/Expert systems are also covered in Chapter 3. More specific guidance on the application of these

data-filling techniques to analogue read-across is given in Chapter 4, and for a broader category approach in Chapter 5. Examples of the data matrices used to report the use of this approach is shown in Chapter 7.

2.3.1 Category membership

In an ideal situation, a category would include all potential members of the category when first developed. This ideal situation will be difficult to achieve in practice. For example, even when a category includes all the single compounds that can be included, it may not necessarily include the additional commercial products that are complex substances containing a mixture of compounds which are also included in the category.

Practical considerations will often influence the choice of chemicals included in the category. Since categories have often been developed in the context of a High Production Volume Chemicals programme, the selection of the chemicals that are included in a particular chemical category has frequently been guided by the fact that the chemicals in the category are produced in high volumes and likely to be dependant on which chemicals are manufactured by the consortium of companies sponsoring the category.

However, it should be noted that the category may also contain substances that are not produced in high volumes, or indeed, substances that are not necessarily commercially available, as well as other substances put on the market by companies not involved in the category evaluation. Substances included in the category that are not formally evaluated have previously been described as "surrogate" substances. This term is not used in the guidance as these substances may subsequently be assessed, e.g. if their production volume changes.

There are significant potential advantages associated with the evaluation of a category which contains a high proportion of its potential members. The conclusions drawn from the evaluation are likely to be more robust, since the category evaluation is less likely to be affected by the subsequent addition of other substances, and the potential advantages of limiting animal and other testing are also likely to be greater.

As chemical categories submitted to authorities for review often do not contain all potential members of a category, due to the practical considerations outlined above, they are evaluated based on the data available for the chemicals submitted. If subsequently chemicals are assessed which fit within the definition and rationale of the category, the category might have to be re-evaluated based on the available data for those additional chemicals.

A substance can potentially belong to more than one category. For example, a multifunctional compound can belong to a category based on function A as well as to the category based on function B. The properties of the compound will be influenced by the presence of both functional groups.

2.3.2 Assessment of categories and individual compounds in a category

The successful use of a category approach should lead to the identification and characterisation (qualitative or quantitative) of the hazards for all the members of the category, irrespective of their production volume or whether or not they are produced by the companies carrying out the category evaluation.

If a substance is assessed and subsequently identified as a member of an existing category, it will be necessary to evaluate both the data for this substance in the light of the category evaluation and the category evaluation in the light of the data for the additional substance. If the initial category evaluation is sufficiently robust, the additional data is unlikely to alter the conclusions of the initial evaluation significantly. Since subsequent assessments of additional members of a category are possible at any time, there is an incentive to ensure that as many potential members of a category are included in the initial evaluation as possible. This would ensure that the evaluation is sufficiently robust in order to minimise the potential revisions as a result of additional data at a later date.

Experience has shown that in many cases additional chemicals are identified which fall on either the lower or upper end of an existing category. In those cases additional testing might be necessary to confirm that the chemicals belong to the category. In these cases, best professional judgement and Weight of Evidence (see section 3.5) are used together in making recommendations/decisions about whether to test or not .

When assessing whether a substance could be a member of an existing category (but it is not already listed as such), the concept of "applicability domain" may be useful. The applicability domain (AD) of a (sub)category would identify the structural requirements and ranges of physicochemical, environmental fate, toxicological or ecotoxicological properties within which reliable estimations can be made for the (sub)category members. For example, there may be a trend of increasing acute aquatic toxicity with increasing chain length from C2 up to a carbon chain length of C12, after which no aquatic toxicity is seen because the water solubility has decreased with increasing chain length. Thus the applicability domain for aquatic toxicity would be C2 to C12.

2.3.3 Subcategories

In some cases, an effect can be present for some but not all members of the category. An example is the glycol ethers, where the lower members of the category show reproductive toxicity whilst higher members do not. In other cases, the category may show a consistent trend where the resulting potencies lead to different classifications. Examples include the lower aliphatic ethers, where aquatic toxicity is insufficient to lead to classification for aquatic toxicity with the lower members of the category, but does lead to classification for this effect with higher members (Hart J & Veith G, 2007).

In these cases it can be helpful to divide the category into subcategories. Examples which have been encountered within the OECD HPV Chemicals Programme (http://cs3-hq.oecd.org/scripts/hpv) include the case of mono-, di-, tri-, tetra-, and penta- ethylene glycols, when a subcategory was denoted by a cut-off of chain length of 6-8 to account for the change in physical form from liquid to solid and a decrease in uptake. A slightly different approach was used in the case of Oxo alcohols C9 to C13 where clear trends in properties were seen with increasing chain length (Caley J et al., 2007). For environmental hazards, two category members exhibited higher ecotoxicity than the other five members and thus formed a subcategory in the assessment. For the long chain alcohols (C6-22 primary aliphatic alcohols), decreasing water solubility and increasing lipophilicity is observed with increasing chain length, leading to a cut-off for acute aquatic toxicity effects at C13 to C14 and around C15 for chronic effects. At C>18, biodegradability is reduced. Three distinctive subcategories can be identified using the GHS classification criteria for aquatic toxicity based on the trends in toxicity and biodegradability.

Subcategories may arise for a number of reasons and are often endpoint specific:

a) an effect which varies in intensity across the category, such that some members of the category meet the criteria for one hazard classification for the particular endpoint, whereas other members of the category meet the criteria for another. These subcategory definitions can be qualitative (i.e. they have degrees of hazard potential or different regulatory classifications) or quantitative (the numerical values of the endpoint include values on either side of a breakpoint).
b) an effect where there is a peak in activity or a breakpoint in a trend can also lead to the formation of subcategories.
c) it is possible that a trend analysis may apply to a subcategory but not to the whole category.

The concept of subcategories has been introduced to improve the practicality and flexibility of the category approach and it does not alter the scientific basis of the category approach.

2.4 The mechanistic basis of chemical categories

A category of chemicals will often show the presence, absence or modulation of a particular effect for all members of the category, based on the presumption of a common mechanism of action. This can be expected to apply to many different categories of chemicals for many aliphatic hydrocarbons, aliphatic amines, nitriles, aldehydes, alcohols, and ethers (Jäckh R, 2007). Additional examples can be found from the OECD HPV Chemicals Programme [http://cs3-hq.oecd.org/scripts/hpv].

If the data for a category includes one or more exceptions to the effects expected from a common mechanism of action, a review of the toxicological data for the category should be able to explain the difference in behaviour. Excluding the exception(s) from the category would decrease the information content of the category and hence its robustness. The presence of such “outlying” effects underlines the importance of developing an understanding of the (toxic) mechanisms of action within categories.

A category may be justified on more than one basis, for example both a chain length and metabolic pathway category (Caley J et al., 2007). Multiple justifications could increase confidence in the category. This increased confidence is largely a result of the more detailed evidence that the common mechanisms of toxic action have been properly identified.

In principle, a category is not endpoint-specific, since the structural changes across the category would be expected to produce changes that would affect the whole spectrum of properties of the individual members in a coherent and consistent manner. The changes in properties across a category, for each parameter, would be the result of related rather than purely arbitrary differences. However, it is recognised that in practice it may be possible to identify the trends and changes for some but not all of the properties of potential interest, and hence it may not be possible to use a category approach to identify all relevant effects.

One example is the use of a metabolic pathway approach where the category approach will be able to address the common toxicological mechanism for endpoints related to systemic effects, whereas it may not predict the local effects (on skin and other membranes) due to the parent compound (see for example the category of monoethylene glycol ethers and their acetates or diethylene glycol ethers and their acetates [http://cs3-hq.oecd.org/scripts/hpv] (Caley J et al., 2007))

For some series of compounds, the lower or upper end of the series may show marked changes in effects. At the lower end of the series, the methyl analogue may have exceptional properties. Examples are the differences shown in acute toxicity between methyl alcohol and ethyl alcohol, and for carcinogenicity

between butter yellow and its ethyl homologue or between methylcarbamate and ethylcarbamate. This may be the result of specific differences in metabolism, such as the differences in carcinogenicity between benzene and toluene, due to the possibility of metabolism of the methyl group with carboxylate formation (Jäckh R, 2007).

The presence of a breakpoint can indicate a change in the mode of action or the effect of a consistent tendency across a category. In a homologous series of organic compounds, there is often a breakpoint e.g. the loss of aquatic toxicity as carbon chain length increases and solubility decreases.

The importance of a common mechanism of action is also a factor in deciding what chemicals would not be expected to be relevant members of a category. Variations in chemical structure can affect both toxicokinetics (uptake and bioavailability) and toxicodynamics (e.g. interactions with receptors and enzymes). For example, the introduction of a carboxylate or sulfate function often decreases bioavailability and toxicity to mammals, whilst halogen substituents tend to increase lipophilicity and increase toxicological activity (see example in Worth A et al., 2007). Thiols and esters are not considered as relevant analogues for evaluation of ether activity (see example in Hart J & Veith G, 2007).

2.5 Application of the chemical category approach

In cases where the approach to chemical hazard and risk assessment is based on the evaluation of substances on an individual basis (e.g. the approach taken for the notification of new substances) testing requirements are primarily based on the production volume of the chemical. This approach is consistent with the fact that the legal obligations are placed on individual producers, and as a result, producers are legitimately concerned to provide information on their own product, but do not necessarily have any interest in acquiring data on related substances in which they have no commercial interest.

As stated in Section 2.2.1, since categories have often been developed in the context of a High Production Volume Chemicals programme, the selection of a particular chemical category has normally been guided by the presence of a number of chemicals in the category that are produced in high volumes. However, it should be noted that a category may also contain other substances that are not HPV chemicals (or indeed, are not necessarily commercially available). These chemicals are still members of the category, and may prove to be relevant candidates for further testing in order to evaluate the properties of the category as a whole.

The formation of a category has in many cases also been dependant on which chemicals are manufactured by the consortium of companies sponsoring the category. Different industry sectors may co-operate on category assessments. This guidance recognises that it is a challenge for Industry to include all relevant members based on the basic properties excluding use pattern/exposure. There may be different needs for hazard information for different members of a consortium depending on uses and thereby the outcome of the risk assessments for the individual members of the chemical category. It is therefore important to develop incentives or articulate benefits for industry taking this approach, as it would be desirable for the consortium to check with other producers/manufacturers for appropriate support and information.

The chemical category approach can be very beneficial when information from other category members help to fill data gaps for untested chemicals. However, the approach may not always be straight forward, especially when a category has many members, when the trend analysis does not show an obvious trend, and/or when different kinds of information (e.g., computational data as well as experimental data) are available within a category. The experience from the OECD HPV Chemicals Programme, where

industry has the opportunity to discuss their category approach with a Sponsor Country, has shown that this collaboration is very helpful. It is therefore recommended that, for "difficult" categories, the assessor should consult the relevant regulatory authorities when developing a category approach. For substances that are part of the OECD HPV Chemicals Programme, the OECD will continue to support collaboration between industry and authorities.

2.6 Robustness of a chemical category

A number of factors contribute to the robustness of a category. Useful considerations might include:

a) membership of the category characterised by the number of members in a category and the available data
b) the density and distribution of the category (both in terms of the chemicals represented and the data available)
c) the quality of the underlying experimental data for each of the endpoints covered
d) the presumed mechanistic basis underpinning the category for a particular endpoint
e) the quality of the data estimated by the external computational approaches

The current document does not provide criteria for validation of chemical categories. Instead the document provides guidance on how to optimise the robustness of chemical categories and how to document the justification for each category.

2.7 The interdependence between categories and QSARs

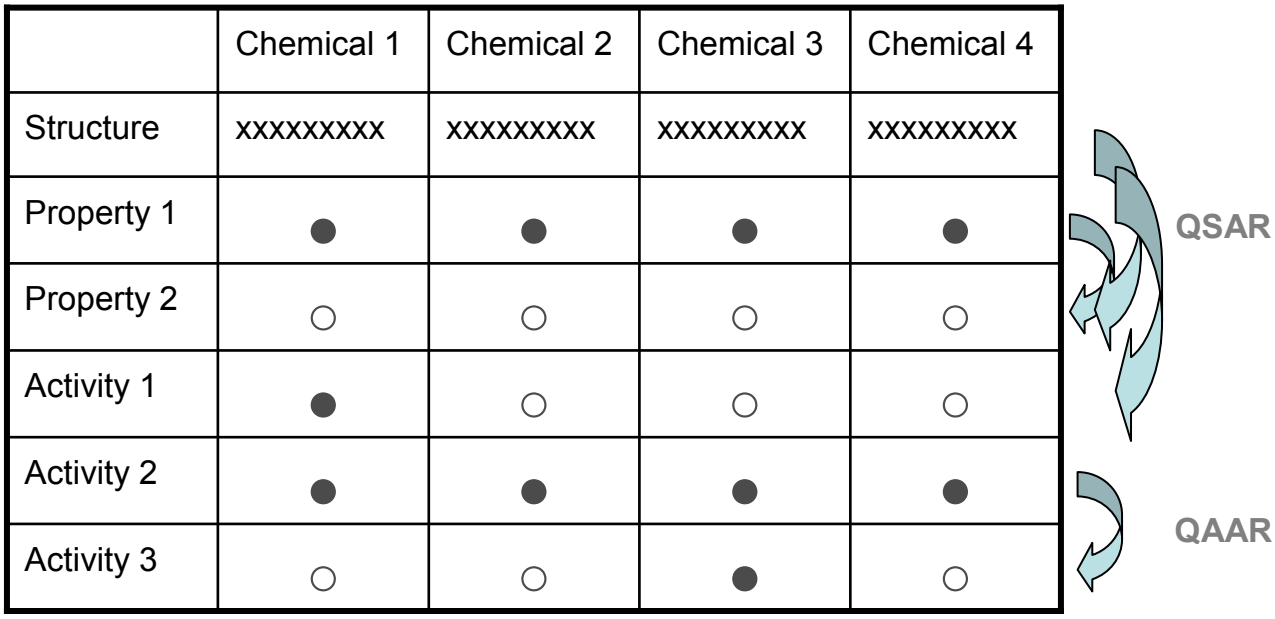
The chemical category and QSAR concepts are strongly connected. The concept of forming chemical categories and then using measured data on a few category members to estimate the missing values for the untested members is a common sense application of QSAR. The reason this concept is so compatible with QSAR is that this broad description of the categories concept and the historical description of QSAR are one and the same (Figure 2).

A Quantitative Structure-Activity Relationship (QSAR) is a quantitative (mathematical) relationship between a numerical measure of chemical structure, and/or a physicochemical property, and an effect/activity (Figure 2). QSARs often take the form of regression equations, and can make predictions of effects/activities that are either on a continuous scale or on a categorical scale. Thus, in the term “QSAR”, the qualifier “quantitative” refers to the nature of the relationship, not the nature of the endpoint being predicted. An example of a QSAR is the prediction of acute toxicity to an invertebrate species (*Tetrahymena pyriformis*) by means of a regression equation with the partitioning behaviour (logKow value) of the chemical as a descriptor (Schultz et al., 2002).

Similarly, a Quantitative Activity-Activity Relationship (QAAR) is a mathematical relationship, but between two biological endpoints (Figure 2), which can be in the same or different species. QAARs are based on the assumption that knowledge about the mechanism or mode of action, obtained for one endpoint, is applicable to the “same” endpoint in a different species, or to a similar endpoint in the same

species, since the main underlying processes are the same (e.g. partitioning, reactivity, enzyme inhibition). QAARs provide a means of performing trend analysis and filling data gaps[3].

Figure 2. Graphical representation of a QSAR/QAAR



Footnote to Figure 2: A QSAR can make extrapolations from chemical structure and/or physicochemical properties to other properties or activities. A QAAR makes an extrapolation from one activity to another related activity.

The common scientific foundation between forming categories and QSARs/QAARs is that chemicals, once grouped together on a basis of common structural attributes, become chemical classes which exhibit consistent trends in their chemical properties and biological hazards. In addition, these trends in chemical activity are often related directly to trends in chemical structure expressed by QSARs.

In many cases, QSARs are quantitative models of key mechanistic processes which result in the measured activity of the chemicals. The importance of this mechanistic understanding is two fold. First, the structure-activity relationships provide useful models for hypothesis testing which increases the reliability and causality of the QSAR model. Secondly, the mechanistic understanding can be described as a series of structural requirements which define the mechanism boundaries on reliable domain of application of QSAR model.

[3] The experience with QAAR is currently limited and therefore this approach has not been routinely used. The concept is presented in this document for completeness sake. Further experience in the application of this concept will lead to revisions of this document.

The categories concept creates a practical and powerful approach for describing these structural requirements of toxicity mechanisms. Chemicals can be grouped together initially using expert judgement which is reflected by the chemicals included. Further discussion may question the similarity of some chemicals based on measured data, evidence of anomalous behaviour or other information about the chemical attributes which suggest some chemicals may fit more than one category. The careful use of expert judgement to define the boundaries of a chemical category is crucial to the reliable application of QSAR models or other methods to estimate values for untested chemicals. A formal definition of which chemicals should be included in a category and which chemicals should be excluded can lead to much more reliable estimates of missing values than the use of QSAR models with poorly defined domains. The expert judgement should be described in a transparent manner in order to be evaluated by others.

A QSAR estimate is the result of an assumption and a prediction about the chemical. The assumption is that of the predominant interaction mechanisms of the chemical, and thus leads to selection of a QSAR model. The prediction is the quantitative estimation of the intensity, or potency, of the chemical structure within the specific mechanisms of interaction. Both the assumption of mechanism and the prediction bear heavily on the reliability of that overall QSAR estimation.

However, the errors created in selecting the proper QSAR model for a specific chemical are greater than those related to the potency estimate of the QSAR model. For example, in ecotoxicity studies, some phenols are polar narcotics, some are uncouplers, and others are electrophilic. QSAR models for each mechanism have comparable uncertainty, but the potency of the latter mechanism can be orders of magnitude greater than polar narcotics. The use of a category approach can thus help to ensure that the QSAR estimates are based on mechanistically valid models by aiding correct selection of the model.

Further information on the use of (internal) QSARs to express trends in categories, and on the use of (external) QSARs to provide additional support for trends, is given in Sections 3.3 and 3.4, respectively.

Within a chemical category, the primary difference between hazard identification and classification and labelling is that the classification and labelling is performed in the context of risk management thresholds established by the regulator. It is possible that the risk management threshold is defined simply as a positive test result in a hazard identification test guideline and the majority of a category would be expected to be classified similarly. However, if the risk management threshold is a specific value along a large range of possible potency values for a specific hazard endpoint, it is reasonable to expect some member to be above or below that threshold and still belong to the chemical category. For classification and labelling, the QSAR models may be designed to either provide a potency estimate or to estimate the likelihood that the potency would be above or below the risk management threshold.

Estimation methods work best for homologous series of chemicals where the metric for extrapolating from one chemical to another is a simple molecular weight, number of carbon atoms or a similar parameter which can be linked to physicochemical properties of the chemicals. However, when the members of the category are not a simple homologous series, it is essential that some parameter which predicts the trend across the members be established in order to extrapolate the measured values to the missing values. For example, the vapour pressure is mechanistically related to the acute inhalational toxicity (LC_{50}) of ethers because it is a surrogate for the thermodynamic activity of the chemical in the blood and tissues (Hart J & Veith G, 2007); but it is not directly related to carbon number or molecular weight because the degree of branching is significantly different among the category members. An estimate using carbon number would not produce defensible extrapolations within this category. In contrast, vapour pressure is a more reliable parameter to extrapolate the results from measured values to missing values.

In addition to the concern over which parameter to use in the estimation, it is necessary to make an assumption about the proportionality factor so that the structural differences between a measured and unmeasured chemical can be proportioned into a difference in toxicity. For example, the acute inhalational toxicity (LC_{50}) of ethers does not increase with vapour pressure with a proportionality of 1.0, but rather with a proportionality of 0.7 (see example taken from Hart J & Veith G, 2007). The advantage of a more rigorous use of QSAR models within categories is that one can base the estimate in the large context of a mechanistic model where the parameter for extrapolation and the proportionality factor(s) are easily justified and explained in transparent terms.

CHAPTER 3. APPROACHES TO DATA GAP FILLING IN CHEMICAL CATEGORIES

3.1 Introduction

The absence of relevant, reliable and sufficient experimental data for a chemical, results in one or more data gaps which need to be filled in order to finalise the hazard and/or risk assessment. This chapter explains the following non-testing techniques for filling data gaps:

a) read-across
b) trend analysis and use of computational methods based on internal models
c) use of computational methods based on external models

In principle, these techniques can be used to indicate either the presence or the absence of an effect. In certain cases , the application of these techniques to assess a particular chemical may benefit from the generation of test data for one or more other chemicals in the category. In other words, the generation of additional experimental data by strategic testing may be useful.

In this document, the term "model" refers to any formalised method for estimating the properties of chemicals, and typically refers to a QSAR, QAAR or expert system. These models are only useful for data gap filling when they are based on data of sufficiently high quality. This is particularly important when applying a model to the interpretation of boundary substances.

The use of these three techniques is described in more detail below. It should however be recognised that whilst these three techniques are described separately in the following section, there are many elements that are common to all three approaches. All three techniques can be used with varying degrees of applicability in the context of both the analogue approach and to a wider category approach. Experience from current practice shows that the first of these three techniques, the use of qualitative or quantitative read-across is already widely used and is often accepted as a valid approach for regulatory purposes. Whilst computational approaches based on SARs, QSARs, QAARs or expert systems can also provide a basis for filling data gaps, experience shows that additional supporting evidence is often required for acceptance of these estimates.

3.2 Read-across

In the read-across technique, endpoint information for one chemical is used to predict the same endpoint for another chemical, which is considered to be "similar" in some way (usually on the basis of structural similarity). In principle, read-across can be applied to characterise physicochemical properties, environmental fate, human health effects and ecotoxicity. For any of these endpoints, read-across may be performed in a qualitative or quantitative manner. In practice, read-across for basic physicochemical properties is not generally recommended, since reliable data should normally be available or easily obtainable, does not involve the use of animals and provides key information for the assessment of a

chemical. However, there may occasionally be practical problems, especially for UVCBs, when the use of these techniques will be required.

Within a group of chemicals, read-across can be performed in the following ways to fill data gaps:

a) one-to-one (one analogue used to make an estimation for a single chemical)
b) many-to-one (two or more analogues used to make an estimation for a single chemical)
c) one-to-many (one analogue used to make estimations for two or more chemicals)
d) many-to-many (two or more analogues used to make estimations for two or more chemicals)

The transition between comparisons using an analogue approach involving more than two chemicals and a more comprehensive category approach described in the following chapter is of course arbitrary. The guidance on read-across given below applies both to the analogue approach described in Chapter 4 as well as to the categories approach described in Chapter 5.

It should be recognised that the robustness of a category approach would be expected to be considerably greater than that of an analogue approach, since the basis for evaluating any individual chemical in the category is greater, and there is usually more measured data available in such a wider approach. The following sections contain guidance particularly with respect to supporting information that is more relevant for the use of an analogue approach, as a category approach will in itself provide additional support for the robustness of the estimates.

A chemical being used to make an estimate can be referred to as a source chemical, whereas a chemical for which an endpoint is being estimated can be referred to as a target chemical.

Read-across can be qualitative or quantitative. In qualitative read-across, the presence (or absence) of a property/activity for the target chemical is inferred from the presence (or absence) of the same property/activity for one or more source chemicals. Qualitative read-across gives a 'yes/no' answer. In quantitative read-across, the known value(s) of a property for one or more source chemicals is used to estimate the unknown value of the same property for the target chemical. Quantitative read-across is used to obtain a quantitative value for an endpoint, such as a dose-response relationship.

Most often, structural similarity and similar properties and/or activities between chemicals is used as a basis for read-across. Thus, endpoint information is read-across from a structural analogue. A structural analogue is a source chemical whose physicochemical and toxicological properties are likely to be similar to the target chemical as a result of structural similarity. The similarity may be based on the following:

a) a common functional group (e.g., aldehyde, epoxide, ester, metal ion). An example is the ethylene glycols category assessed in the OECD HPV Chemicals Programme (http://cs3-hq.oecd.org/scripts/hpv),
b) a common precursor and/or breakdown product, that results via physical or biological processes (metabolic pathway similarity). This is used to examine related chemicals, such as acid/ester/salt. Examples are certain azo dyes based on carcinogenic components such as benzidine or other carcinogenic aromatic amines, where the carcinogenic aromatic amine is formed by the metabolism of the dye.

Analogies between chemicals can also be drawn on the basis of common mechanisms of action and similarities in chemical (or biochemical) reactivity.

In principle, it is possible to predict the presence or absence of a property/effect by applying the read-across approach. Read-across from a negative result is regarded as equally valid and convincing as a positive result provided the test design, concentrations tested etc. have been chosen adequately. For example, if all tested chemicals of a category are shown not to be mutagenic and if there is scientific justification that the untested chemical rightly belongs in the category, it is justified to assume that the untested chemical is also not mutagenic. However, if the mutagenicity test system that has been used is inappropriate to demonstrate the genotoxicity of the group of chemicals, then a conclusion that the category would not be mutagenic would not be valid. There is extensive experience of read-across of negative findings or absence of effect in the EU risk assessment and classification and labelling work and the OECD HPV Chemicals Programme. For example, in the assessment of medium-chain chlorinated paraffins (both within the EU Existing Substances Regulation and the OECD HPV Chemicals Programme), data from the short-chain chlorinated paraffins was used as supporting evidence for lack of genotoxicity, low acute dermal toxicity and absence of skin sensitisation potential. It is particularly important to adequately justify read-across of negative findings. The read-across approach is most robust when a quantitative trend between the analogues can be established.

A stepwise approach for performing read-across on a limited number of chemicals (analogue approach) is given in Chapter 4. The use of this approach for filling data gaps in a larger category approach is shown in Chapter 5.

3.2.1 Choice of qualitative or quantitative read-across

Before deciding on the type of read-across approach which is necessary, it is important to determine why the data gap is being filled and what type of data is required. Is a specific value required or does the endpoint need to be checked against a threshold or hazard banding/cut-off (for example a classification banding)? Read-across has been used for a range of different reasons to date, for example:

- To fill a data gap for a specific endpoint - both threshold and non-threshold values[4]
- To reduce an assessment factor used to derive a PNEC[5]
- To flag a concern for further testing[6]
- To read-across classification and labelling[7]

In deciding on whether to use quantitative or qualitative read-across, the nature of the property should also be considered. It may be expressed on a numerical or categorical scale. In most cases, a

[4] For example, the assessment of short chain chlorinated paraffins CAS 85535-84-8 where the NOAEL for effects via lactation was read-across from medium chain chlorinated paraffins (both within the EU Existing Substances Regulation and the OECD HPV Chemicals Programme, http://cs3-hq.oecd.org/scripts/hpv/).

[5] For example, the assessment of medium chain chlorinated paraffins CAS 85535-85-9 where aquatic toxicity data from short chain chlorinated paraffins was used to show invertebrates are most sensitive and thus reduce the assessment factor from 50 to 10 to derive the PNECaquatic (both within the EU Existing Substances Regulation and the OECD HPV Chemicals Programme, http://cs3-hq.oecd.org/scripts/hpv/).

[6] For example, the assessment of p-t-butylphenol CAS 98-54-4 within the EU Existing Substances Regulation) where data from p-t-pentylphenol were used to request further testing on endocrine disruption in fish (Tsakovska I & Worth A., 2007).

[7] For example, the common EU classifications for skin irritation and sensitisation agreed for sulphate, dichloride, nitrate and carbonate salts of nickel (Hart J., 2007).

specific value is required for risk assessment, such as a NOEC or NOAEL, environmental half-life or partition coefficient. A numerical value obtained by quantitative read-across would normally be needed. For conducting a hazard assessment, PBT assessment or assigning classification and labelling, one generally needs to know whether that substance fits the particular hazard criteria. Identification of the hazard by qualitative read-across may be adequate.

An issue that may arise when read-across is carried out in the context of a category is that the experimental results for different category members may be available for different test methods or species relating to the same general endpoint. For example, in the case of reproductive toxicity, only screening studies may be available for some category members, whereas two-generation studies may be available for other members. As the estimated results from the category approach have to be useful for risk assessment and classification, the uncertainty associated with the underlying results has to be ascertained. It is clear that the scope of the estimated results for a member of a category cannot exceed the scope of the underlying data for the other members of the category, e.g. if for genotoxicity, only in vitro results are available for some members of the category (source chemicals), only conclusions on in vitro genotoxicity can be reached for the members of the category for which experimental results are lacking (target chemical). If the scope of the underlying experimental results for an endpoint vary (e.g. a mix of results from screening tests and higher tier tests), it is necessary to clarify the scope of the estimated results for the category members for which no experimental results are available. It may be possible to apply a weight-of-evidence approach to all the data, which could lead to the same hazard identification for all the members of the category, irrespective of the data available for the individual compounds.

3.2.2 Qualitative read-across

In qualitative read-across, the presence or absence of a property is inferred from the established properties of one or more analogues. The main application of qualitative read-across is in hazard identification, and usually results in the allocation of the target chemical(s) to the same hazard category as the source chemical(s).

The arguments to support the read-across are normally based on expert (eco)toxicological judgement. Several factors can be considered in making this judgement. The assumption that a common substructure is responsible for the common property or effect could be affected by interactions between the substructure and other parts of the chemical structure. Another substructure could alter the property/effect in a qualitative manner (in which case the assumption may be false) or a quantitative manner (i.e. change the degree to which the substance exhibits the property). One example could be changes in the degree of branching of a carbon chain which can affect biodegradability and toxicity. In addition to interactions between substructures, differences in one or more whole-molecule properties could alter the assumption of commonality (e.g. differences in aqueous solubility could affect the read-across of a classification for aquatic toxicity). These factors are assessed by a process of expert judgement. However, it should be recognised that expert judgement may not necessarily be accepted by all concerned in the evaluation. An example is the read-across of carcinogenicity for musk ketone, which was evaluated by the SCHER (2006).

If a regulatory classification is used to express the property or effect, a quantitative change in the potency of the chemical could be sufficient to warrant a different classification, depending on the classification threshold. If a difference in the potency between source and target chemicals is suspected, for example based on trends in the available data, a quantitative read-across approach rather than a qualitative approach would usually be required. This is particularly important where the target chemical is suspected to have a more stringent classification than the source chemical. A different classification can be considered where the classification criteria are based on the strength of the available evidence rather than a

quantitative cut-off. In addition, differences between a direct and an indirect effect can lead to a different classification of the target chemical than the source chemical. An example is the classification of benzidine azodyes as category 2 carcinogens whilst benzidine itself is classified as a Category 1 carcinogen.

3.2.3 Quantitative read-across

In addition to identifying a particular property for a target chemical, in quantitative read-across the known value of a property for the source chemical(s) is also used to estimate the unknown value of the same property for the target chemical.

When applying quantitative read-across, there are four general ways of estimating the missing data point:

a) by using the endpoint value of a source chemical, e.g. the closest analogue in a (sub)category[8]
b) by using an internal QSAR (see Section 3.3) to scale the available experimental results from two or more source chemicals to the target chemical[9]
c) by processing the endpoint values from two or more source chemicals (e.g. by averaging, by taking the most representative value)
d) by taking the most conservative value of the closest analogues or the most conservative value in the (sub)category[10]

Quantitative read-across can also be utilised for complex substances/UVCBs, typically by applying data from physicochemically similar substances (e.g. substances with similar boiling ranges, carbon ranges, composition) or by applying data from key/major constituents. However, this must be done carefully, may be more applicable for indication of ranges and requires an understanding of the key structures that may drive the behaviour of UVCBs. This is further discussed in section 6.5.

In risk assessment, a dose descriptor is used as a quantitative basis for deriving a Predicted No Effect Concentration (PNEC) or Derived No Effect Level (DNEL), depending on the endpoint. To account for various sources of uncertainty in the derivation of the PNEC or DNEL, an assessment factor is applied to a numerical value of the dose descriptor.

When conducting a risk assessment, a NOAEL, NOEC or other effect concentration such as EC10 may be read-across in order to derive a DNEL or PNEC for the target chemical, provided that this is justified. Read-across of the PNEC or DNEL itself from the source to target chemical is not recommended since the range of available data for a chemical must be considered when deriving the DNEL or PNEC. The size of the assessment factor used to derive a PNEC or DNEL depends on the confidence with which it

[8] For example, the OECD HPV Gluconates category, where aquatic toxicity data for Sodium D-gluconate were read-across to the calcium and potassium salts, D-Gluconic acid and Glucono-delta-lactone (Caley J et al., 2007).

[9] For example, OECD HPV C6-22 Aliphatic Alcohols category where internal QSARs were developed to predict aquatic toxicity based on Kow and thus derive aquatic toxicities for the target chemicals (http://cs3-hq.oecd.org/scripts/hpv/).

[10] For example, the assessment within the EU Existing Substances Regulation and the OECD HPV Chemicals Programme of Zinc distearate used aquatic toxicity data from the more soluble zinc salts (chloride, sulphate) to derive the PNECaquatic for Zinc distearate (Tsakovska I & Worth A., 2007).

can be derived from the available data. Generally, lower assessment factors can be used with larger more relevant datasets.

When deriving a DNEL or PNEC based on an endpoint which has been read-across, it is important to ensure that the read across is sound and that the target chemical is unlikely to be more potent than the source chemical. In cases where there are multiple source chemicals, and consequently a range of possible values for read-across, the use of the most conservative (lowest) value may be sufficient to account for the uncertainty in the read-across. In particular, the read-across is likely to be conservative when the target chemical has a lower bioavailability than the source chemical. If there is any uncertainty in the read-across, and thus the DNEL or PNEC derived from it, it may be necessary to conduct testing for that endpoint.

In the assessment of medium chain chlorinated paraffins CAS 85535-85-9 (both within the EU Existing Substances Regulation and the OECD HPV Chemicals Programme, http://cs3-hq.oecd.org/scripts/hpv/), aquatic toxicity data from short chain chlorinated paraffins was used to show that invertebrates are most sensitive and thus reduce the assessment factor from 50 to 10 to derive the PNECaquatic despite the fact that no chronic fish test was available for medium chain chlorinated paraffins.

There is no experience to date with the use of DNELs for human health risk assessment so further guidance should be developed on the use of read-across data in DNEL derivation once experience is gained with its use.

In cases where there are concerns that the relative potency of the different chemicals may be sufficiently large to affect the conclusions of either hazard identification (in cases where the criteria contain a quantitative cut-off) or risk assessment (based on an estimated PNEC/DNEL), additional testing specifically designed to demonstrate differences in potency across a category can be considered.

3.2.4 Choice of endpoints for the application of read-across

In principle, read-across can be applied for any property or endpoint, irrespective of whether it is a physicochemical property, environmental fate parameter, human health effect, or ecotoxicological effect.

In practice, read-across is not encouraged for basic physicochemical properties (e.g. water solubility, logKow) since these properties provide key information for the assessment of a chemical in particular for the assessment of the environmental properties, and experimental data or valid QSAR predictions should normally be available (or should be reasonably obtainable).

3.2.5 General considerations when performing read-across

Irrespective of the type of read-across, it is important to consider a number of factors (Hanway & Evans, 2000):

a) Whether the data point of the source chemical is relevant and reliable for the purpose of the read-across. If read-across data have not been produced using the most current OECD test methods, particularly careful consideration of the quality and suitability of a method is important.
b) Whether the source and/or target chemical is a multi-functional compound and whether the additional functionality may therefore affect the reliability of the read-across.

c) The purity and impurity profiles of the target and source chemicals need to be assessed. There is a need to identify those impurities which might influence the overall toxicity of the source chemicals and to discuss the consequences these impurities will have for the the robustness of the chemical category and hence for the read-across. If all category members have the same sort of impurities, then they may not have any relevant influence on the read-across. If there is a very biologically active impurity (e.g. CMR substances) in one category member, but not the other members, then the results from that category member might not be appropriate for read-across.
d) Comparison of the physicochemical properties of the target and source chemicals, particularly the physical form, molecular weight, water solubility, particle size and structure[11], partition coefficient and vapour pressure, provides useful information as to their similarity.
e) The likely toxicokinetics of the substances, including the possibility of different metabolic pathways coming into play, needs to be considered where possible.
f) Information from valid (Q)SARs may be used where possible to inform decisions on the need, extent and type of additional testing.

In the case of UVCBs (Section 6.5), it should be considered whether the differences between the UVCBs in a specific group would actually give rise to different effects, bearing in mind the internal consistency of the basic structural families and assumption of similarity of action or reaction.

3.2.6 Supporting information

It is important to provide supporting information to strengthen the rationale for the read-across. Thus, in addition to the property/endpoint being read-across, it is also useful to show that additional properties, relevant to the endpoint, are also (qualitatively or quantitatively) similar between the source and target chemicals. Such properties could be known or suspected determinants of the endpoint, or they could be limiting factors.

Relevant molecular properties of the source chemical should be of comparable value to those of the target chemical. The selection of relevant molecular properties depends on the endpoint for which the read-across is being performed. The identification of these properties could be based on expert knowledge, or could be based on the use of properties (molecular descriptors) that have been found to be useful predictors of the endpoint in QSAR models.

In the case of single substances, irrespective of the endpoint being read-across, useful considerations might include:

a) the presence or absence of additional functional groups or substituents that could influence the behaviour of a chemical
b) similarity in physicochemical profiles (e.g. MW, logKow, water solubility)
c) similarity in other toxicological and/or ecotoxicological data
d) the likely toxicokinetics of the substances, including the possibility of different metabolic pathways coming into play, needs to be considered where possible.
e) information from valid (Q)SARs may be used where possible to inform decisions on the need, extent and type of additional testing.

[11] There is debate ongoing on the regulatory application (classification and derivation of dose-descriptors).

In cases where there are convincing arguments for a read-across approach, the need to generate new data with tests on vertebrates should require a strong and convincing argument, whether to remove an unwanted classification or confirm a non-classification. In such cases, if test data demonstrate the measured value differed considerably from the estimated, the read-across and the resultant category, if applicable would have to be carefully reconsidered. A weight of evidence analysis (section 3.5) may be useful for determining whether the read-across or the test data was suspect.

In the case of UVCBs (Section 6.5), it should be considered whether the differences between the UVCBs in a specific group would actually give rise to different effects, bearing in mind the internal consistency of the basic structural families and assumption of similarity of action or reaction.

3.2.7 Supporting information for environmental endpoints

What constitutes appropriate supporting information will depend on the environmental endpoint being read-across. However, basic physicochemical properties that determine environmental distribution and fate (e.g. MW, partition coefficients such as logKow, water solubility) will generally be useful. Particle size and structure[11] may also be relevant.

For example, in the case of aquatic toxicity, similar logKow and aqueous solubility values between the source and target chemicals could be used to support the read-across, because logKow is known to be a determinant of the toxicity in aquatic organisms when the effect is mediated by mechanisms of narcosis. If the chemical is known or expected to act by a non-narcotic mode of action, additional properties might provide useful supporting information. For example, experience with new chemicals in the EU suggests that tests such as acute toxicity to Daphnia can provide additional confidence that read-across of other data is possible, i.e. if toxicity differences are found between the source and target chemical then further testing for other endpoints may be appropriate (Hanway & Evans, 2000). The acute Daphnia toxicity test raises few animal welfare issues while providing good confirmation of the comparability of aquatic toxicity.

Furthermore, in the case of read-across of aquatic toxicity endpoints, results (fish, invertebrates and algae) for source and target chemicals should be compared. For example if a read-across to acute toxicity to fish is based on a presumed mode of action, and if this mode of action is applicable to invertebrates and algae, the available results for invertebrates and algae for the source and target chemicals should confirm the applicability of the read-across.

3.2.8 Supporting information for human health endpoints

What constitutes appropriate supporting information will depend on the human health endpoint being read-across. However, physicochemical properties that determine biokinetics and bioavailability (e.g. MW, partition coefficients such as logKow, water solubility, pH, vapour pressure, viscosity) will generally be useful. Particle size and structure[11] may also be relevant.

In general, current practice relies heavily on expert judgement. The type and amount of supporting evidence needed may vary with the endpoint concerned.

In the case of musk ketone, the target chemical, read-across for carcinogenicity can be based on the data for musk xylene, the source chemical (SCHER, 2006). Important considerations for the read-across were:

a) musk ketone (the target chemical) has similar physicochemical properties as musk xylene (the source chemical)
b) there are structural similarities between the two chemicals
c) both chemicals have been tested for mutagenicity; neither chemical is genotoxic
d) both nitro musks are inducers of cytochrome P4502B1
e) However, musk xylene effects on the liver cytochrome P450 activities are different from those of musk ketone. While both musk xylene and musk ketone induce CYP 2B gene expression, the induced cytochrome P450 2B protein is present in an inactivated form after musk xylene administration resulting in a much lower CYP 2B1 associated catalytic activity. Due to its chemical structure, musk ketone cannot be reduced to an enzyme inhibiting p-amino metabolite and therefore induces, but does not inactivate CYP 2B enzymes in mice. Hence, high levels of active cytochrome P450 2B are present after administration of musk ketone.
f) The mode-of-action of musk xylene in both mice and rats seems to be identical, while some species differences in the pattern of cytochrome P450 induction by musk ketone are observed
g) The role of enzyme induction in the development of liver tumours by musk xylene in mice and in the toxicity of repeated administration of musk ketone is not well defined.
h) There are similarities of the effects of both musk xylene and musk ketone to effects of phenobarbital, which also induces liver tumours in rodents by a non-genotoxic mode-of-action and is also an inducer of cytochrome P450 2B.
i) Assuming that the induction of cytochrome P450 2B is a relevant mode-of-action for liver tumours induction by musk xylene, read-across based on "enzyme induction" and structural and physicochemical properties may be sufficient as a basis for read-across since musk ketone is also an inducer of this enzyme. More detailed information on the mechanisms of enzyme induction by musk ketone is not available.

For some endpoints, such as skin sensitisation or mutagenicity, chemical reactivity might provide useful supporting information. For skin sensitisation, one of the necessary hurdles a chemical has to undergo is to form a stable association with a skin protein. This is thought to be a covalent association where the chemical behaves as an electrophile and the protein as a nucleophile. A similar analogy is relevant for mutagenicity but where DNA represents the nucleophile. An experimental system that quantifies the electrophilic reactivity would be useful to support a read-across for skin sensitisation, (Aptula et al., 2006) or mutagenicity (Benigni et al., 2005).

In vitro data might also provide useful supporting information. For example, if acute mammalian toxicity is being read-across, it might be appropriate to refer to similarity of *in vitro* cytotoxicities of the source and target chemicals, if it is known (or suspected) that cytotoxic effects underlie the acute systemic effect. Relationships between *in vitro* cytotoxic effects and acute systemic toxicity has been investigated by a number of workers (e.g. Clemedson et al., 2002).

3.3 Trend analysis and computational methods based on internal models

For a given category endpoint, the category members are often related by a trend (e.g. increasing, decreasing or constant). The trend could be related to molecular mass, carbon chain length, or to some other physicochemical property. For larger categories, it is possible that several different relationships can be established for a single endpoint, thereby defining subcategories. A chemical that identifies a turning point in a trend is called a breakpoint chemical (see also Section 2.2.3). Category members falling at the opposite extremes of a trend and within which interpolations are considered reliable are called sentinel (boundary) chemicals.

A demonstration of consistent trends in the behaviour of a group of chemicals is one of the desirable attributes of a chemical category and one of the indicators that a common mechanism for all chemicals is involved. When some chemicals in a category have measured values and a consistent trend is observed, missing values can be estimated by simple scaling from the measured values to fill in the data gaps.

The observation of a trend (increasing, decreasing or constant) in the experimental data for a given endpoint across chemicals can be used as the basis for interpolation and possibly also extrapolation (see Figure 1). Interpolation is the estimation of a value for a member using measured values from other members on "both sides" of that member within the defined category spectrum, whereas extrapolation refers to the estimation of a value for a member that is near or at the category boundary using measured values from internal category members. Interpolation between measured analogues may give a more reliable result depending on the reliability of the measured data. Interpolation can be performed when the series of values is monotonic (all increasing or decreasing) or when non-monotonic (e.g. parabolic). In such circumstances the extent to which the available data describe the trend will determine the level of confidence in the prediction.

In general, interpolation between category members is preferred to extrapolation. However, it may be the case that whilst data is available for several members of a category, there can be data gaps for the boundary chemical. In this case extrapolation will be necessary. It should be noted that extrapolation based on a clearly established trend will be substantially more robust than the use of read-across from analogues to fill a data gap. The robustness of any extrapolations used to fill data gaps will be closely related to the general evaluation of the whole category.

When establishing trends in data, laboratory and experimental variations should be considered. Similar species/strains, endpoints and test protocols should be compared. Deviations from a trend should be clearly identified and possible reasons for the deviations laid out in the category analysis.

In principle, it is possible to predict the presence or absence of a property/effect by applying trend analysis. The category approach is most robust when a quantitative trend between the category members can be established. A lack of observed toxic effects for a chemical substance in a study of a specific endpoint (especially if no dose-relationship can be established because no effects are observed at some of the doses tested) requires further consideration and. in such circumstances, the data need to be carefully evaluated. It is important to distinguish between cases where the lack of response can be explained on the basis of the mechanistic understanding for that endpoint, or whether the tests have failed to demonstrate the absence of an effect for the category as a whole.

The larger the category, the more likely that there may be breaks in trends which may affect the reliability of interpolation or extrapolation. The observation of a "break" in a trend among some members of a category is a warning sign, but is not necessarily an indication that the chemicals with different trends exhibit different toxicity pathways. Bioassay measurements frequently are only comparable over a narrow range of chemical properties with the result that different pharmacodynamic factors are controlling the bioassay results for different chemicals. The bilinear or multilinear nature of trends in measured data, if observed, can be used to confine the methods for scaling intensity of the endpoint to specific members of the category.

The observation of a trend "break" should not be confused with differences in the hazard classification of the members of a category. When the cut-off dividing different classification bands is between the extreme values of the trend, then the members of the category will be classified differently. If all members of the category have properties above or below the administrative cut-off agreed for that

property, the trend analysis may be useful for judging the adequacy of forming the category but apparent breaks in the trends would not lead to differences in the classification.

There is little current experience in the use of the type of formal trend analysis shown here. However, there is good reason to believe that arguments based on this approach would be acceptable to estimate missing data, and that this technique provides a basis for a robust estimate.

The data for a particular endpoint can be used to construct a QSAR that describes the properties of the members of the category. A Quantitative Structure-Activity Relationship (QSAR) is a quantitative (mathematical) relationship between a numerical measure of chemical structure, or a physicochemical property, and an effect/activity. QSARs often take the form of regression equations, and can make predictions of effects/activities that are either on a continuous scale or on a categorical scale. Thus, in the term "QSAR", the qualifier "quantitative" refers to the nature of the relationship, not the nature of the endpoint being predicted.

An example of a QSAR is the prediction of acute toxicity to an invertebrate species (*Tetrahymena pyriformis*) by means of a regression equation with the partitioning behaviour (logKow value) of the chemical as a descriptor (Schultz et al., 2002).

A trend might also be expressed as a quantitative activity-activity relationship (QAAR). A Quantitative Activity-Activity Relationship (QAAR) is a mathematical relationship between two biological endpoints, which can be in the same or different species. QAARs are based on the assumption that knowledge about the mechanism or mode of action, obtained for one endpoint, is applicable to the "same" endpoint in a different species, or to a similar endpoint in the same species, since the main underlying processes are the same (e.g. partitioning, reactivity, enzyme inhibition).

Thus, a chemical category can be seen as a set of "internal" QSARs (and possibly also internal QAARs) for the different endpoints, with the advantage that all the underlying data are transparently available to the assessor. Such models provide quantitative descriptions of the trends within a category and are referred to as "internal" QSARs (or QAARs) because they are derived directly from the experimental data for the category members. These models are also likely to be "local" models in the sense that they are based on a relatively small data set. Such an internal local model was for example developed for acute aquatic toxicity for the category of long-chain alcohols (C6-22 primary aliphatic alcohols) assessed within the OECD HPV Chemicals Programme (http://cs3-hq.oecd.org/scripts/hpv).

Such methods work best for homologous series of chemicals where the metric for extrapolating from one chemical to another is a simple molecular weight, number of carbon atoms or a similar parameter which can be linked to physicochemical properties of the chemicals. However, when the members of the category are not a simple homologous series, it is essential that some parameter which predicts the trend across the members be established in order to extrapolate the measured values to the missing values. For example, the vapour pressure is mechanistically related to the acute inhalational toxicity (LC_{50}) of ethers (Hart J, 2007) because it is a surrogate for the thermodynamic activity of the chemical in the blood and tissues; but it is not directly related to carbon number or molecular weight because the degree of branching is significantly different among the category members. An approach using carbon number would not produce defensible extrapolations within this category. In contrast, vapour pressure is a more reliable parameter to extrapolate the results from measured values to missing values.

3.4 Computational methods based on external models

In this guidance document, the term "external model" is used in distinction to the "internal model" described in the section above and can refer to any model (QSAR, QAAR or expert system) that was not developed as part of the category formation process. If such models are used to fill data gaps in a category, they should be based on experimental data that are obtained from a wider range of chemicals than those used in the category. Such external models are also known as "global models" since the data on which they are based comes from a relatively large number of chemicals in comparison with those in the category. In this sense, the category under evaluation is a subcategory of this wider QSAR.

The predictions made by an external model may be used to provide additional support for the trend (even though reliance is usually placed on the experimental data rather than the model estimates). To be applicable the prediction should be considered as reliable and the comparison between the predicted value and the experimental value available for other members of the category or the analogue should be taken into account. For example, a parabolic QSAR could be used to characterise the trend in bioconcentration factor (BCF) values across a series of chemicals of increasing molecular weight.

In other cases, model predictions may be used to identify and rationalise category members that deviate from a trend. For example, a QSAR or expert system might indicate that certain chemicals in a series have anomalous behaviour due to metabolism, although this would need to be confirmed by consideration of the biological plausibility of the differences.

If multiple experimental data are available for a single substance, the result of a computational model can be helpful in choosing a valid data point.

The result of one or more computational models can be used to increase the confidence in an experimental measurement for a single substance. For example, within the EU Existing Substances Regulation, estimated results obtained with two QSAR models for biodegradation were used to support an experimental observation of ready biodegradability for acrylaldehyde (Tsakovska I & Worth A 2007).

3.5 Weight-of-evidence considerations

Since the data used in a hazard assessment should be relevant, reliable and sufficient for the regulatory purpose, it is necessary to base the assessment on the totality of available information, i.e. to apply weight-of-evidence (WoE) considerations. The WoE assessment can be based on experimental data as well as estimated data (obtained by applying one or more non-testing approaches). In most cases, estimated data might be used to supplement and increase confidence in the available experimental data, whereas in some others, such data might be used instead of experimental data.

Further guidance on WoE considerations is provided in the OECD Manual for Investigation of HPV Chemicals (OECD, 2007b)

CHAPTER 4. GUIDANCE ON A STEPWISE PROCEDURE TO PERFORM THE ANALOGUE APPROACH

4.1 Introduction

This chapter provides guidance on how to estimate missing data from a single or limited number of compounds using the analogue approach.

The guidance in this chapter is primarily based on the widespread current experience in the application of read-across using the analogue approach using non-formalised approaches. However, the guidance also provides indications of where computer-based methods can be included to facilitate the process. A stepwise approach to analogue evaluation is proposed, in which the use of formalised computational approaches can be integrated.

In the EU, there is considerable experience in the application of read-across using the analogue approach in the classification and labelling group (ECB, 2005, Comber M & Simpson B, 2007, Gallegos Saliner A et al., 2007, Hart J 2007, Hart J & Veith GD, 2007, Schoeters I & Verougstraete V, 2007). More recently additional experience has been gained in the risk assessment of Existing Chemicals (ESR programme; (Tsakovska I & Worth A, 2007), and in the Notification of New Substances (NONS programme; Hanway & Evans, 2000).

There is also considerable experience on the use of analogue approaches in the OECD HPV programme and by the US EPA (ECB, 2005). Within the OECD HPV Chemicals Programme, read-across has been extensively performed since 1998. Examples of initial hazard assessments that rely on data from analogues, and which have been published, include: isobutanol (CAS No 78-83-1), p-chlorotoluene (CAS No 106-43-4), and methyltriacetoxysilane (CAS No 4253-34-3). These initial assessments are available from UNEP Chemicals (2006).

Much of this experience has taken place in the context of consultation in either the EU Technical Committees or at the OECD, and reflects a consensus on the use of expert judgement between experts from the member countries.

The current practice in the EU is often based on an empirical identification of an appropriate analogue. The choice of analogue is normally fairly straightforward, as any potential analogue has to be data-rich in order to form a basis for comparison. In many cases the choice is governed by the availability of data on an analogue manufactured by the same producer or an analogue where data are available from detailed regulatory evaluations (OECD HPV Chemicals Programme or the EU Existing Substances Programme) or from the open literature. For example, under the EU Existing Substances Programme, data for ETBE was estimated by comparison with the data collected for MTBE and TAME (Tsakovska I & Worth A., 2007).

It is foreseen that read-across using the analogue approach using non-formalised methods will continue to be the more frequently used method for filling data gaps over the next few years. Based on a learning-by-doing approach, the experience gained in application of this approach will lead to further improvements of this guidance in the future.

In the case of single substances, or complex substances where there are dominating constituents, read-across by non-formalised approaches generally involves the identification of a chemical substructure that is common to the target chemical and its analogue (or their respective breakdown products) and the assumption that:

a) in the case of qualitative read-across, the presence (or absence) of a property/activity for the chemical of interest (target chemical) can be inferred from the presence (or absence) of the same property/activity for the analogue (source chemical).

b) in the case of quantitative read-across, the known value of a property for the analogue (source chemical) can be used to estimate the unknown value of the same property for the chemical on interest (target chemical). In the case of a toxicological effect (human health or ecotoxicological), this assumption implies that the potency of an effect shared by the two chemicals is similar or follow a regular pattern.

In the case of complex substances, the basis for comparison is likely to be different. For example, complex substances derived from certain process streams may share common structures.

With limited information it can be difficult to judge the degree of uncertainty associated with the assumption of commonality for a particular read-across. To provide the most robust read-across possible, other relevant properties should be compared between the source and target chemicals.

4.2 Stepwise approach to read-across using the analogue approach

The following stepwise approach is recommended, but should be regarded as flexible and not the only possible approach, see Figure 3.

4.2.1 Step 1: Identification of potential analogues

There are a number of different possible ways of identifying potential analogues as source chemicals with data with which the target chemical can be compared.

In many cases, the choice of a source chemical is straightforward. Similar chemicals produced for similar uses by the same company (or sector group of companies) are often used as potential analogues. In this case, no formal selection techniques are used.

However, a more formal search strategy may indicate additional potential analogues for comparison, and hence, increase the robustness of the read-across. It should be noted that with increasing numbers of chemicals included in a read-across, the closer this approach is to the approach used for categories described in the next chapter. One starting point would therefore be to consider whether the chemical is best evaluated by an analogue approach, or whether a wider category approach should be used. One factor that would affect the choice is whether the chemical is a member of a category that has already

been evaluated. Another factor would be the number of analogues identified: if a significant number of analogues are identified, then a wider category approach would be justified, as outlined in the next chapter.

Information on categories that have been evaluated by the US EPA is available from http://www.epa.gov/opptintr/newchems/pubs/chemcat.htm

Information on categories that have been evaluated within the OECD HPV Chemicals Programme is available from http://cs3-hq.oecd.org/scripts/hpv/

There is no single information source on categories evaluated within the EU. However, information can be found in ECB, 2005, Gallegos Saliner et al. (2007) and Tsakvoska & Worth (2007).

A number of industry sectors have applied the principles of "grouping" for use in evaluation of health and environmental hazard properties. Examples, including rationales for grouping, include petroleum substances (Concawe, 2001), dyes and pigments (ETAD, 2001), chlorinated paraffins (CPIA, undated), surfactants (CESIO, 2000, 2003) hydrocarbon solvents (HSPA, 2002), acrylate resins (UV/EB Acrylate Resins, 2003), petroleum additives (ATC, 2000a, b) and bitumen (Eurobitume, 2002) (see ECB, 2005).

Categorisation approaches have been applied to flavours and fragrances (Salvito D, 2007) under JECFA, USHPV, Environment and Health Canada DSL Program, SPORT, and the safety assessment of fragrance ingredients under RIFM.

Computational methods for analogue selections are expert knowledge in combination with electronic substructure searching and automatic tools using molecular similarity indexes (e.g. the Tanimoto similarity index). The pharmaceutical industry, which is the predominant user of the concept of molecular similarity, is employing similarity methods in a wide range of applications e.g. virtual screening, estimation of absorption, distribution, metabolism, excretion and toxicity (ADME/Tox) and prediction of physicochemical properties (solubility, portioning, etc.). Whilst these techniques have not been widely used in this context, the use of such techniques should be considered when searching for relevant source chemicals for comparison.

A non-exhaustive list of possible analogue-searching tools is given in Table 1.

The identification strategy is an exploratory process, and is not intended to be an element of the read-across rationale. If a large number of analogues are identified, the use of the categories approach described in the next Chapter is recommended. It should also be noted that the use of a category approach reduces the demands on extensive data for any individual source chemical, as this approach draws on the cumulative data available for all the individual chemicals in the category.

The structural similarity and the purity and impurity profiles of the substance and the structural analogue need to be assessed. The fundamental basis for any read-across decision must be that the chemical structures of the analogues are sufficiently close for there to be a reasonable expectation of similar effects. The more divergent the structures, the lower will be confidence in making such a prediction. In general, where biologically active functional groups are present, they should be present in both structures and be in the same structural orientation so that any biological activity would be unaffected.

The extent to which differences in the purity or impurities are likely to influence the overall toxicity (Hanway RH, 2000), needs to be addressed and, where technically possible, excluded (see also 3.2.5 point (c)).

4.2.2 Step 2: Data gathering for the analogues

For the source analogues chosen, published and unpublished data should be gathered on standard physicochemical properties, environmental fate and transport properties, ecotoxicological and toxicological effects. Standard physicochemical properties include physical state, MW, logKow and other partition coefficients (e.g. the Henry's Law coefficient, soil organic-carbon partition coefficient), aqueous solubility, particle size and structure[11], vapour pressure, melting point and boiling point. Since these physicochemical properties provide basic information on environmental distribution, fate and bioavailability, they can often provide supporting information for the read-across. The data gathering should include all existing relevant data, including both experimental data and data generated by non-testing methods.

If a large number of analogues are identified, it is recommended to consider forming a larger chemical category (see chapter 5). If this is not feasible, e.g. for practical reasons, computational tools, such as (Q)SARs can help to reduce the dataset to a subset of the closest analogues, e.g. homologues for which properties similar to the target chemical are estimated (see Sections 3.2.6 and 3.2.7).

Data are already available on many high volume chemicals that have been thoroughly assessed. Information on substances assessed by the OECD is available from the OECD (http://cs3-hq.oecd.org/scripts/hpv) and the United Nations: (UNEP Chemicals, 2006).

Information on chemicals assessed in the EU can be found on the ECB website (http://ecb.jrc.it).

Information on the environmental and human health effects of chemicals can be found from a large number of internet-accessible databases. A list of such databases, including internet links, has been compiled by the European Chemicals Bureau (http://ecb.jrc.it/QSAR/information_sources/information_databases.php).

4.2.3 Step 3: Evaluation of available data for adequacy

Where data are available from relevant peer-reviewed sources such as the OECD HPV Chemicals programme, EU risk assessment programme or other comparable sources, the data can normally be used without further evaluation.

In other cases, the available experimental data should be evaluated for adequacy e.g. using the OECD Guidance for Determining the Quality of Data for the SIDS Dossier (see section 3.1 of the *OECD Manual for Investigation of HPV Chemicals,* OECD, 2007b).

If read-across data have not been produced using the most current test methods, particularly careful consideration of the quality and suitability of a method is important (Hanway & Evans, 2000).

4.2.4 Step 4: Construct a matrix of data availability

A matrix of data availability should be constructed for the target endpoint and all other relevant endpoints (see Appendix 1 for an example). The matrix should include the chemical of interest (target chemical) and the analogue(s) (source chemical(s)). If multiple analogues are identified, they should be arranged in a suitable order (e.g. according to molecular weight). The ordering should reflect a trend or progression within the group. The cells of the matrix should indicate whether data are available or unavailable. If possible, the cells should also indicate the available reliable key study results.

4.2.5 Step 5: Assess the adequacy of the analogue approach and fill the data gap

It is currently only possible to provide limited guidance about how to decide whether data from an analogue can be used to fill a data gap, and the decision remains largely an expert judgment. Similarly, it is not possible to provide definite guidance on how data gaps could be filled quantitatively by read-across.

However, the factors shown in section 3.2.5 need to be addressed when evaluating the results of a read-across using an analogue approach. The supporting evidence discussed in sections 3.2.6, 3.2.7 and 3.2.8 should also be considered.

Wherever possible, the relevance of the read-across of other endpoints should be evaluated in the light of the known or suspected mode of action. The applicability of the read-across can also be evaluated in the light of available data for both source and target chemical for other endpoints where the mode of action is likely to be similar. The use of QSAR predictions can also be useful to assess the applicability of the read-across, both by predicting the missing data and comparing the experimental data available and the predictions.

Chemicals that cannot be represented by a molecular formula or structure can be handled on a case-by-case basis, depending on the components of the complex substance and on the data available for the complex substance and/or components.

If the read-across is considered to be suitable, the missing data for the target chemical(s) is evaluated using the data from the source chemical(s) according to the guidance in Chapter 3.

If the read-across is not considered to be suitable, three options are possible. It may be necessary to identify alternative analogues – the best analogues may indeed not have the relevant experimental data, so it may be necessary to choose analogues of lower quality in order to obtain data - or the use of a more extended category approach can be considered. It may also be necessary to obtain the information directly by testing.

4.2.6 Step 6: Document the analogue approach

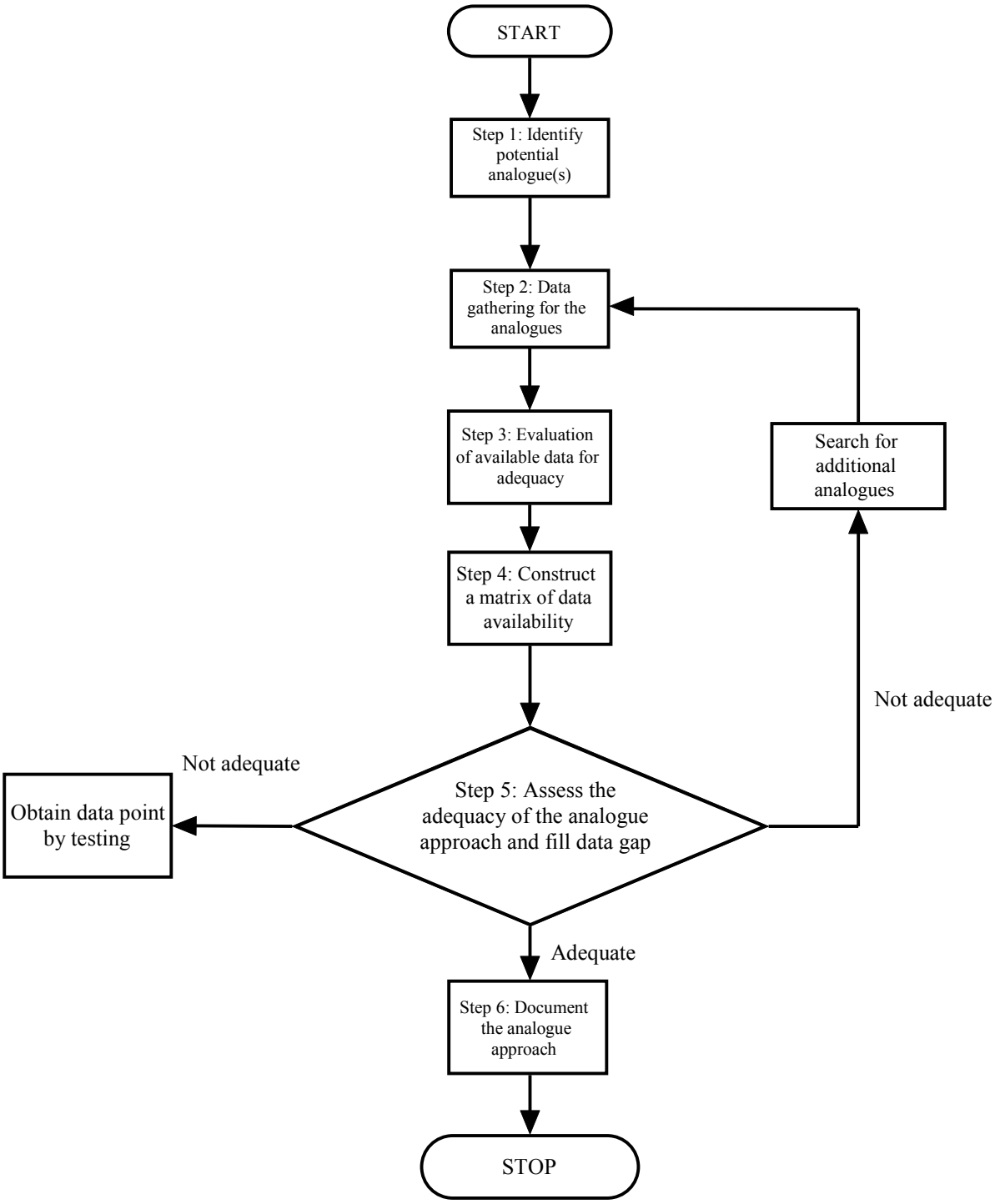
If the read-across is considered to be suitable, the approach should be documented according to an appropriate format in order to justify that the approach may be used instead of testing (see Chapter 7). The justification for the read-across should include an explanation of the rationale, as well as the assessment including all relevant supporting information. Ideally examples of unsuitable read-across should also be documented.

Table 1. Selected tools for analogue-searching

Tool & Website	Remarks
OECD (Q)SAR Application Toolbox www.oecd.org/env/existingchemicals/qsar	Toolbox containing a library of (Q)SAR methods and databases of experimental results, as well as tools to form chemical categories and fill data gaps by read-across, trend analysis and (Q)SARs. Proof-of-concept version to be publicly available in March 2008 Contains ca. 200,000 records. Searchable by chemical name, CAS number, SMILES, substructures, mechanisms of reaction etc.
AIM	US EPA's Analog Identification Methodology. Links to publicly available, experimental toxicity data for target chemical as well as structural analogues Due to be publicly available in 2007. Contains 31,031 records. Searchable by CAS number, SMILES and (sub)structure.
Ambit http://ambit.acad.bg	Chemical databases and functional tools, including a tool for defining applicability domain of QSAR models Developed by IdeaConsult Ltd Publicly available Contains 463,426 records. Searchable by chemical name, CAS number, SMILES and (sub)structure.
ChemFinder http://www.chemfinder.com	Publicly available and subscription scientific databases. Searchable by diverse parameters including chemical name, synonyms, CAS number, formula, chemical structure (exact match, substructure, similarity search), toxicological and physico-chemical properties.
ChemID Plus http://chem.sis.nlm.nih.gov/chemidplus	Publicly available database from the US National Library of Medicine (NLM). Contains over 379,000 records. Searchable by chemical name and CAS number.
Hazardous Substances Database (HSDB) http://toxnet.nlm.nih.gov	Publicly available toxicology database on the National Library of Medicine's (NLM) Toxicology Data Network (TOXNET) More than 4800 peer-reviewed records. Searchable by chemical name, fragment name, CAS number, subject terms.
Danish (Q)SAR Database	Publicly available version of the QSAR database developed by

http://ecbqsar.jrc.it	DK EPA, and made available by ECB website. Contains 166,000 records. Searchable by chemical name, CAS number, endpoint, and (sub)structure
Leadscope http://www.leadscope.com	Commercially available databases and (Q)SAR functionalities Searchable by chemical name, (sub)structure, toxic effect, study type, and experimental conditions.
SciFinder http://www.cas.org/SCIFINDER	Commercially available and internet-accessible portal to extensive collection of chemical and biochemical information from scientific literature and patents. Searchable by chemical name, (sub)structure, biological sequence and reaction, as well as by research topic, author, and company.

Figure 3. Stepwise approach to an analogue approach



CHAPTER 5. GENERAL GUIDANCE ON A STEPWISE PROCEDURE TO DEVELOP CATEGORIES

5.1 Introduction

Chemical categories accomplish the goal of obtaining hazard information through the evaluation of all available experimental data for the individual chemicals in the category, so that reliable estimates that are adequate for classification and labelling and/or risk assessment can be made without further testing of the individual members of the category. If there are sufficient experimental data to support the category evaluation that the chemicals in the category behave in a similar or predictable manner, then the relational features described in Figure 1 can be used to assess the chemicals instead of conducting additional testing. If not, it may be necessary to: a) perform limited and targeted testing; b) revise the category hypothesis (and therefore the applicability of the category in terms of members and/or endpoints); or c) as a last resort abandon the category hypothesis.

The review of the use of chemical categories carried out in preparation for the development of this guidance[12] concluded that the main lessons learned with the use of the chemical category concept are:

a) Initial hazard assessments were agreed upon by OECD member countries for 240 chemicals in 42 different categories as of 2006, by applying the chemical category approach. The approach can therefore be considered to be widely accepted for regulatory purposes.
b) Currently more than a third of the substances assessed yearly within the OECD HPV Chemicals Programme are assessed through the use of chemical categories and this fraction is estimated to increase significantly over the next few years as experience grows in member countries.
c) As already concluded for the US HPV Challenge Programme, chemical categories can be used to estimate results for both environmental and human health endpoints.

The guidance in this Chapter documents a stepwise approach to the formation of categories. The current practice is based on the use of non-computational methods. However, guidance is also included on where computational tools could be used at various steps in this process to support the development of categories. It is emphasised that such computational tools can supplement but do not replace the need for expert judgement, which is required throughout the process. Whilst the use of these tools is considered to be helpful in a category approach, it should be recognised that the use of approaches for which there is little or no regulatory precedence should be used in close collaboration with the relevant regulatory authority.

This chapter should be read with the understanding that the formation of categories can be carried out using the expertise routinely used in hazard identification and risk assessment. However, given the large number and diversity of chemicals that exist, and the extensive number of categories that may be formed, guidance on how to develop and evaluate chemical categories can not be captured in terms of rigid

[12] Modified from ECB, (2005)

rules. Rather this section describes how information on chemical properties and activities and when available, metabolism and mechanisms of action should be gathered and combined with expert judgement to form robust and well rationalised categories, as well as guidance on how to document the justification for each category. Based on a learning-by-doing approach, the experience gained in application of this approach will lead to further improvements of this guidance in the future.

5.2 Stepwise approach to the formation of chemical categories

In order to use the results from a category, it is necessary to demonstrate that a chemical category is robust, and to do this, certain types of information should be documented. In order to collect this information in a systematic and transparent manner, it is recommended to follow a stepwise approach (Figure 4). The general scheme should be regarded as flexible, since there may be alternative ways of most efficiently obtaining the information.

One reason for needing flexibility is that there can be different starting points in category formation. For example, it may be desirable to start from a single chemical, or small group of chemicals, and to identify analogues to establish a larger category. Alternatively, it may be desirable to start from a defined set of chemicals (e.g. a set list of already classified substances), and to find ways of grouping them and finding additional analogues relating to them.

5.2.1 Step 0: Check whether the chemical is a member of an existing category

Before considering whether to develop a category for a group of substances, the first step should be to determine whether the chemical(s) is (are) a named member of a category that has already been evaluated. Information sources on existing categories include:

a) US EPA: http://www.epa.gov/opptintr/newchems/pubs/chemcat.htm
b) OECD: http://cs3-hq.oecd.org/scripts/hpv
c) United Nations: http://www.chem.unep.ch/irptc/sids/OECDSIDS/sidspub.html

A number of industry sectors have applied the principles of "grouping" for use in evaluation of health and environmental hazard properties. Examples, including rationales for grouping, include petroleum substances (Concawe 2001), dyes and pigments (ETAD, 2001), chlorinated paraffins (CPIA, undated), surfactants (CESIO, 2000, 2003) hydrocarbon solvents (HSPA, 2002), acrylate resins (UV/EB Acrylate Resins, 2003), petroleum additives (ATC, 2000a, b) and bitumen (Eurobitume, 2002) (see ECB, 2005).

Categorisation approaches have been applied to flavours and fragrances (Salvito D, 2007) under JECFA, the US HPV Challenge Program, the Environment and Health Canada DSL Program, SPORT, and the safety assessment of fragrance ingredients under RIFM.

If the chemical is a member of a category that has already been evaluated, its inclusion into the new category should be justified. It is usually sufficient to refer to the evaluation of the category when assessing the chemical, and to refer to the results that have been agreed for the category, taking account of the position of the chemical in the category. Where new data are available for some endpoints, these may be used to verify the existing category and could, depending on the results, lead to a revision of the category.

In some cases, a relevant category may exist, but where the chemical of interest has not been specifically included in the category. For example, this can be the case where a category including only a number of HPV chemicals has been evaluated. In this case, it would be appropriate to extend the membership of the currently defined category to include the chemical of interest. For further guidance on the consequences of extending a category in this way see Chapter 2 (Section 2.2).

5.2.2 Step 1: Develop category hypothesis and definition and identify category members

The first step in developing a category is to develop a basis for the proposed grouping of chemicals.

The category definition should list all of the substances and endpoints covered. Chemical category definitions have referred to chemical classes with a common functional group (e.g. epoxides) or chemicals with an incremental and constant change across the category (e.g. a chain-length category).

Although the chemical structure is usually the starting point, a category definition could also refer to a group of chemicals related by a mechanism of action (e.g. non-polar narcotics) or a particular property. In practice, this particular property is largely related to the chemical structure. For example, in the case of hydrocarbon solvents, products were separated into categories based on basic hydrocarbon structure - aliphatic or aromatic - and then further separated based on boiling ranges, carbon number, and other properties. In some cases, the aliphatic hydrocarbon categories were further separated into subcategories based on specific aliphatic structure such as normal or branched aliphatics (IHSC, 2004/2005).

Some categories have been defined in terms of a metabolic pathway, i.e. they have a stepwise metabolic pathway producing the different members within the category with each metabolic step. More detailed examples of how these types of categories have been evaluated are shown in Chapter 6.

In addition, the category definition should describe the molecular structure a chemical must have to be included in the category, including criteria such as carbon chain length, functionality, and chemical or metabolite equivalence considerations.

It is possible to develop and propose a category for a specific endpoint, or a selection of endpoints, rather than for all of the endpoints required for the substance in question, although this restriction should only be applied where strictly necessary. In particular, all the endpoints that can be expected to be relevant for the category should be included. Since a category is based on an underlying hypothesis of a common mechanism of action, the wider the range of endpoints covered, the more robust the results that are obtained from the category approach.

The category hypothesis should also address:

a) the chemical similarities (analogies) and trends in properties and/or activities that collectively generate an association between the members. These features can be regarded as the parameters that hold the category members together.
b) the specific instances of read-across and trend analysis (interpolations and extrapolations), and any specific computational methods that have been used
c) the set of inclusion and/or exclusion rules that identify the ranges of values within which reliable estimations can be made for category members for the given endpoint. These rules, can be

described as the applicability domain for an endpoint and provide a means of extending the category membership to chemicals not explicitly included in the current definition of a category.

Depending on the basis for the category, the individual members of the category are identified.

In many cases, this is done on an empirical and non-systematic basis. In the OECD HPV Chemicals Programme and the EU Existing Substances Programme, chemicals have frequently been grouped on the basis of their obvious structural similarities (e.g. phthalate esters, groups of oil-derived complex substances, metal compounds).

Since categories have often been developed in the context of a High Production Volume Chemicals programme, the selection of the chemicals that are included in a particular chemical category has normally been guided by the fact that the chemicals in the category are produced in high volumes. However, it should be noted that a category may also contain substances that are not produced in high volumes (or indeed, substances that are not necessarily commercially available) and which may have been tested and provide a source of data for the category. These chemicals are also legitimate members of the category, and may in some cases prove to be relevant candidates for testing in order to evaluate the properties of the category as a whole.

The formation of a category has in many cases also been dependant on which chemicals are manufactured by the consortium of companies sponsoring the category. However, it should be noted that a category may also contain substances that are produced by a number of different companies. It is therefore important for industries wishing to use this approach to consider the formation of a consortium (e.g. based on an Industry sector group) in order to obtain appropriate support and information.

However, when developing a category, the possibility of including additional chemicals that had not been initially selected since they did not meet these pragmatic criteria should be seriously considered. Data may be available for these chemicals that can help in the assessment of the target chemicals . Inclusion of these chemicals will increase the robustness of the category, and reduce the possibility that the addition of these chemicals at a future date would lead to revision of the conclusions for the chemicals specifically under evaluation.

There are many approaches to making a list of category members from the use of simple manual approaches to the use of automated computer-based analogue searching methods.

In preparing a comprehensive list of ethers to form a category of low molecular weight ethers with carbon numbers from 2 to 6, permutations of the SMILES notation for these compounds was used (see Hart J & Veith G, 2007). This approach has the advantage of speed and simplicity, but there are also disadvantages associated with the approach. Systematic use of the SMILES notation can ensure that all possible members of a category are included, and the systematic names of the individual members can be derived from the structures. However, it is often difficult to identify the CAS numbers of the substances without additional work. The production process may also vary across the range of a category, leading to the formation of commercial products of varying complexity, and potentially differing impurity profiles, depending on carbon number. Whilst most of the low carbon number ethers are produced as single compounds, many of the higher carbon number ethers are produced as complex substances with varying components. These commercial compounds may have their own separate CAS numbers, and the available data may only be available for the commercially produced complex substance, rather than for the individual compounds identified on the basis of their structure.

In the case of new category proposals, computational methods can help to develop the category hypothesis (rationale) and to define the category in terms of its endpoints and members. The choice of computational method(s) is likely to depend on the starting point of the investigation. For example, the user may start from a single chemical or a small group of chemicals, with the intention of building up a category by drawing on data from multiple sources (bottom-up or systematic approach). Examples of tools that might help include expert systems such as Derek (LHASA Ltd, UK) or other tools such as Leadscope (Leadscope Inc, USA) or AIM (US EPA). In addition, combinatorial methods exist for identifying, *a priori*, the possible permutations of the substituents on a given substructure. Examples of tools capable of this include TSAR or Cerius2. A variety of computer-based analogue-searching tools have been summarised in Table 1 in Chapter 4. In some cases, these techniques may identify compounds which contain more than one isomer, which can give rise to difficulties in estimating the properties of the individual components (see example in Worth A et al., 2007). However, regulatory experience with the use of these computational tools is still limited and further guidance will need to be developed in the near future.

In identifying a category, it is important that all potential category members are described as comprehensively as possible. For potential members of a category, all relevant CAS numbers should be selected. For some substances, there may be more than one CAS number, and studies may contain relevant data reported under different CAS numbers. Due to historic reporting errors, a CAS number used to describe a substance may not accurately describe the substance as marketed. The CAS numbers of members of the category should also be checked against different chemical inventories (e.g. TSCA, EU, Customs Inventories) as these inventories may indicate which CAS numbers are used for marketing the substances and hence for which CAS numbers additional data might be available.

It is important that information on the purity and impurity profiles of all potential category members is collected at the same time as details of the molecular structure. Differing purity or impurities could influence the overall toxicity. For example, a category member may contain a particularly toxic impurity that is not present in the other substances making it difficult or impossible to draw conclusions on the toxicity of other substances in the category. It is therefore important that category members have similar purity profiles or, where they differ, the effect of the differing purity profiles is known.

5.2.3 Step 2: Gather data for each category member

For each member of the category, published and unpublished data should be gathered on physicochemical property(ies), environmental fate parameter(s), toxicological (human health) and ecotoxicity (environmental species) effect(s). This should include all existing relevant data and not be limited to the endpoints that are mandatory within a given programme (e.g. metabolism and cancer studies are relevant but not part of SIDS in the OECD HPV Chemicals Programme). In some cases where estimated data have been included in an internationally accepted evaluation, these estimates can be included on the same basis as other data that have been critically evaluated.

The computational methods described in Step 2 (Chapter 4) can also be used to identify analogues (and corresponding data) that are included in one or more databases. Having identified a range of possible chemicals, one or more databases could then be searched to identify those chemicals for which data are available. Guidance on data gathering for analogues is also given in Section 4.2.2.

Dossiers should be prepared for each category member. Specific guidance on how to prepare Dossiers for chemical categories with the IUCLID software will be developed and made available in a separate guidance document. Reporting formats are described in Chapter 7.

5.2.4 Step 3: Evaluate available data for adequacy

Available data should be evaluated for its adequacy using e.g. the OECD Guidance for Determining the Quality of Data for the SIDS Dossier (see section 3.1 of the OECD Manual for Investigation of HPV Chemicals).

In evaluating the available data for a category, a number of additional factors will apply that are not relevant when evaluating test results for individual compounds.

- Different types of data may be available for the same endpoint. It is clear that the scope of the estimated results for a member of a category cannot exceed the scope of the underlying data for the other members of the category, e.g. if for genotoxicity, only in vitro results are available for some members of the category (source chemicals), only conclusions on in vitro genotoxicity can be reached for the members of the category for which experimental results are lacking (target chemical). If the scope of the underlying experimental results for an endpoint vary (e.g. a mix of results from screening tests and higher tier tests), it is necessary to clarify the scope of the estimated results for the category members for which no experimental results are available. It may be possible to apply a weight-of-evidence approach to all the data, which could lead to the same hazard identification for all the members of the category, irrespective of the data available for the individual compounds.
- An effect that is defined by a particular numerical cut-off may lead to different conclusions for individual compounds. This type of data should be studied carefully to ensure that the compounds are evaluated in a way that reflects the underlying trends across a category. For instance, a series of compounds may give rise to data that shows a borderline positive irritant effect for some members of the category and a borderline negative effect for others. The data should be carefully evaluated to decide whether (a) this reflects accurately a trend across the whole category or whether (b) the uncertainties in the experimental data justify allocating the compounds to different subcategories (in this example, classifying some category members as irritant and not classifying others). If the second option is considered as the most biologically plausible explanation, the conclusion of the evaluation will lead in some cases to a different conclusion than that based on a simple evaluation of the data taken in isolation. Hence, a borderline positive effect can be interpreted as a negative effect in the light of evidence from other compounds in the category. Similarly, a borderline negative effect can be interpreted as positive taking into account the data from the whole category.
- Where the data suggests possible breakpoints, the data should be evaluated to ensure that these reflect a genuine change in properties or effects and are not due to comparison of results from testing carried out in different laboratories, at different times, with different animal strains, etc.
- The data set may contain an apparent outlier, i.e. one category member where there are experimental data that shows the presence of an effect not seen in other category members. This difference can be real, and provide evidence of special conditions relevant to the particular substance (e.g. the chronic and reproductive toxicity of hexane compared to other lower alkanes). Such results need to be evaluated with particular care to establish whether the result reflects a real difference in a mechanism of action across the category or whether the test result should be questioned.

5.2.5 Step 4: Construct a matrix of data availability

A matrix of data availability (category endpoints *vs.* members) should be constructed with the category members arranged in a suitable order (e.g. according to molecular weight). The ordering of the members should reflect any trends or progression seen within the category. The cells of the matrix should

indicate whether data are available or unavailable. If possible, the cells should also indicate the available reliable key study results (see Appendix 1 for an example).

5.2.6 Step 5: Perform a preliminary evaluation of the category and fill data gaps

A preliminary assessment of the category should be carried out to determine whether:

a) the category rationale is supported, i.e. the category does in fact exhibit one or more of trends postulated in Step 1; and
b) the category is sufficiently robust (i.e. contains sufficient, relevant and reliable information on the category members) for the assessment purpose.

This assessment should be carried out for each endpoint, as the category rationale may lead to a relevant assessment for some endpoints and not for others.

This assessment is largely a matter of expert judgement. Assessment of the category rationale and robustness of the category for the particular regulatory purpose is closely related to the approach chosen for filling data gaps for any particular endpoint, and here the guidance in Chapter 3 for analogue read-across, trend analysis and the use of external QSARs should be taken into account.

If the initial assessment indicates that both criteria are satisfied for a particular endpoint, the data gaps can be filled according to the guidance in Chapter 3 and the chemical category can be finalised and documented.

In applying these techniques, the background for the basis on which the category is formed should be reflected in the way techniques are chosen and applied. Hence for some effects, where the test data suggest a uniform property across a group, read-across from the existing data would normally be considered appropriate. In other cases, where there is a trend in aquatic toxicity related to a change in logKow and based on a narcotic mechanism of action, the data gaps may be filled by data from a valid QSAR for the category. Alternatively, the category can be sub-divided into a number of subcategories defined by the breakpoints in the category, and members evaluated within each subcategory.

If the initial category does not satisfy both of these criteria, the following options should be considered:

a) If further examination of the data suggests that there is a pattern of effects for a limited number of chemicals in the group, then the analysis might suggest that the category should be modified e.g. divided into subcategories (return to step 1).
b) If adequate data do not exist, but the structure-based category is reliable for one or more endpoints, then a category approach may still be proposed for these endpoints. Testing of some chemical category members for some endpoints would still be necessary (go to Step 6). The choice of chemicals and endpoints for testing should be scientifically motivated, but is also likely to involve animal welfare and financial considerations, especially in the case of more "expensive" endpoints.
c) If there are adequate data for a given endpoint, but no apparent pattern, the proposed category may not be appropriate and so testing may be required for all remaining category members for that endpoint (i.e. the category is abandoned).

5.2.7 Step 6: Perform and/or propose testing

If the preliminary assessment supports the category rationale (i.e. a pattern or trend is observed), but the category does not appear to contain sufficient, relevant and reliable information to assess all category members, it may be necessary to perform or propose testing.

In proposing additional testing, a number of factors should be taken into consideration.

- Since a category may contain compounds of different production volumes, the standard information requirements (e.g. those stipulated in the *OECD Manual for Investigation of HPV Chemicals* for the OECD HPV Chemicals Programme) may vary from compound to compound within a category. However, there may be strong scientific reasons that the recommended testing should be conducted on lower tonnage category member(s) in order to identify the actual hazards of the category. In which case the test plans should be confirmed with the appropriate regulatory authority.
- The choice of test will be influenced by the results of the preliminary evaluation of the category.
- If there are no data for any of the members of a category for a particular endpoint, full testing of a limited number of carefully selected category members may be considered appropriate.
- When data are already available indicating the presence or absence of a particular effect, tests may be chosen to provide evidence that compounds selected for testing show the effects that have been predicted from the trend of the property. Hence, for a substance in a category where e.g. skin irritation is predicted, a simple *in vitro* test would be sufficient to provide confirmation of the effect.

Test plans for chemical categories should include a category definition, rationale, and matrix of data availability and be accompanied by the Dossiers for each category member.

The rationale supporting a category definition should be as simple and transparent as possible, and should explain why the existing data and proposed testing data allow interpolation or extrapolation to other members of the category that have no data or proposed testing. The category rationale should be documented in the Category Reporting Format, as described in Chapter 7.

The test plan needs to summarise the adequacy of the existing data, and how the proposed testing will adequately characterise the category.

The matrix of data is a useful part of the test plan and provides a tool for consideration and presentation of the available data. The endpoints are rows in the matrix. If toxicity is expected to vary in a regular pattern from one end of the range of category members to the other end (e.g. high toxicity to low toxicity), samples chosen for testing should bracket both ends of toxicity. If the category is large, testing also needs to be performed and/or data should be available for one or more member(s) in the middle of the range of toxicity. Any change in a tendency for a property should be accompanied by data in the adjacent cells in order to define the limits for the resulting subsets of the category or subcategories. Assuming the columns are the category members, there are no rules for the number of columns and cells that must be filled nor the number that can be empty. Acceptability of the matrix will depend on the number of members in the category, the endpoint, and the confidence in the interpolation and extrapolation.

When selecting a sample to test, it should be representative of the substance marketed, including the presence of any manufacturing impurities.

It should be noted that the category test plan is intended to provide information about the properties of the group as a whole rather than the properties of any specific, individual compound. A category test plan may thus identify as key substances for testing substances of little or no commercial importance. Whilst in some cases this may even require the synthesis of chemicals specifically for this purpose, the approach may still prove more economical, both in terms of expense and numbers of animals used for testing, than a more conventional testing strategy based on individual commercially available chemicals.

5.2.8 Step 7: Perform a further assessment of the category

If new test data are generated, the category should be revised and further assessment to determine whether the criteria outlined in Step 5 are satisfied and therefore whether the category can be finalised and documented.

If the results support the category, the testing phase is complete and the chemical category can be finalised and documented. Remaining data gaps can be filled according to the guidance in Chapter 3.

If the results do not support the category, further testing may be carried out, members of the category may be changed (e.g. dividing the category as appropriate), or the category proposal may be dropped altogether. The latter implies that testing will then be done to fill all appropriate endpoints for each category member.

5.2.9 Step 8: Document the finalised category

The finalised category should be documented in the form of a suitable reporting format (see Chapter 7 for proposed format).

Chemicals that cannot be represented by a molecular formula or structure can be handled on a case-by-case, depending on the components of the substance and on the data available for the substance and/or components.

While a category may be regarded as finalised, it may be revised subsequently in the light of new data and/or experience. For example, the category could be extended by including additional chemicals, or may even be redefined by withdrawing one or more substances.

5.3 IT tools for elaborating dossiers for members of chemical categories

IT tools to build dossiers for members of chemical categories and to document the chemical categories have been developed, e.g. IUCLID 5 or HPVIS.

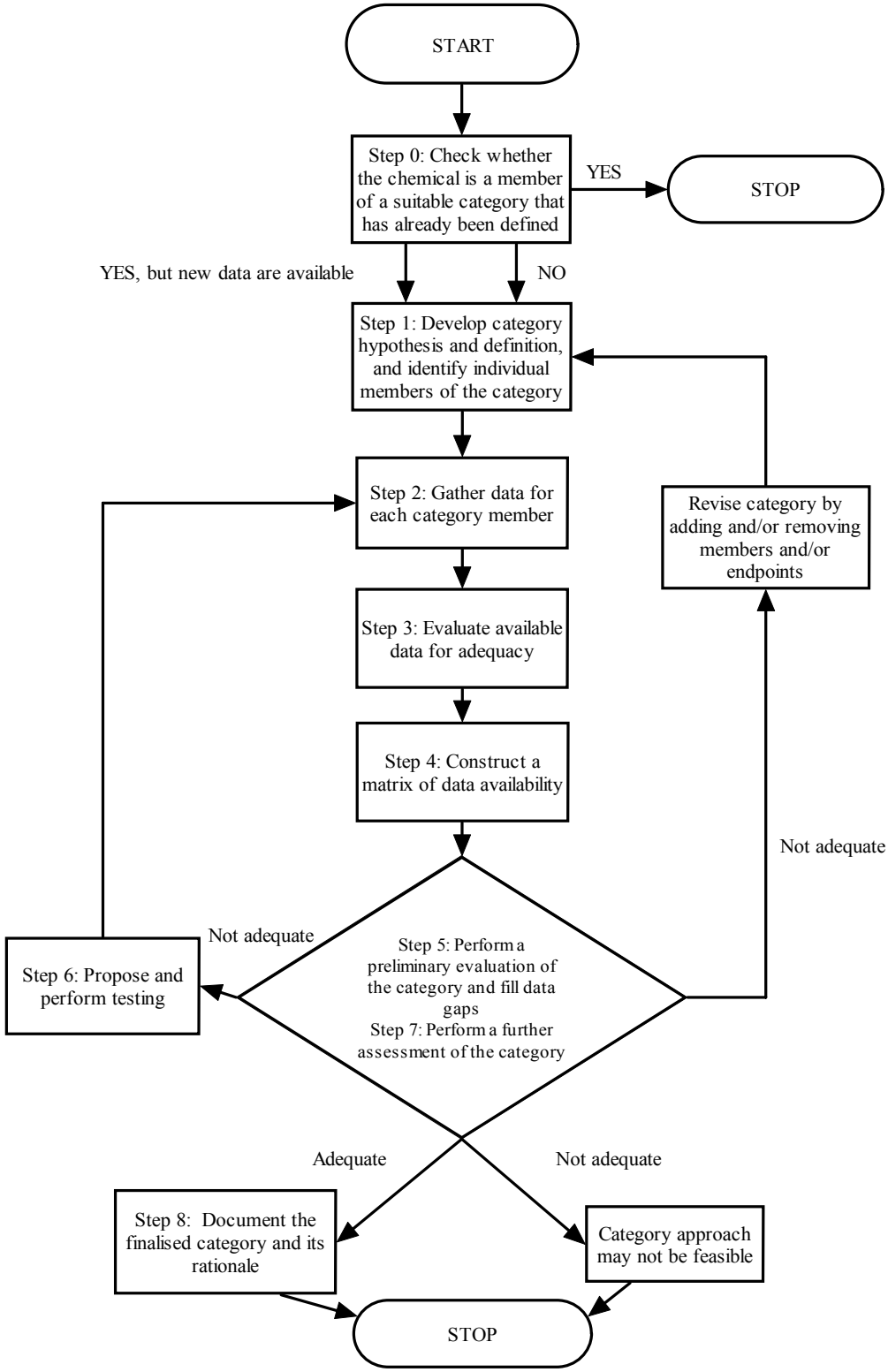
HPVIS has been developed by the US-EPA in the context of the US HPV Challenge Programme.

IUCLID 5 is the recommended tool for submission of dossiers under REACH as well as under the OECD HPV Chemicals Programme.

Both tools, while focusing on the elaboration of dossiers for single substances, allow for the grouping of substances, either for simple analogues or into more complex chemical categories.

Guidance on how to prepare documentation for chemical categories according to the present guidance document with the above mentioned IT tools will be prepared separately.

Figure 4. Stepwise approach to category development



CHAPTER 6. GUIDANCE ON SPECIFIC TYPES OF CATEGORIES

In this chapter, guidance is provided for some specific types of chemical categories. It should be highlighted that the categories described in this chapter are not the only category types that might ever be formed or created.

6.1 Chain length

Chain-length categories show an incremental, and usually constant, increase in chain length across the category. It is assumed that each category member exhibits the same toxic mode of action unless there is a good scientifically demonstrated reason to believe this is not the case. Examples include the homologous series of alpha-olefins, where each category member differs by a methylene group (–CH_2– unit), and the ethylene glycols, where there is an incremental increase in the number of CH_2CH_2O groups. Examples of chain length categories which have been assessed within the OECD HPV Chemicals Programme are alpha-olefins (CAS Nos: 592-41-6, 111-66-0, 872-05-9, 112-41-4, 1120-36-1), higher olefins (CAS Nos: 25264-93-1, 25339-56-4, 25377-83-7, 27215-95-8, 25339-53-1, 25378-22-7, 85535-87-1, 629-73-2, 112-88-9) or monoethylene glycol ethers (CAS Nos: 2807-30-9, 111-76-2, 112-25-4) (UNEP Chemicals, 2006).

Categories defined by chain length generally show an incremental change in molecular weight and other physicochemical properties, such as water solubility or logKow. However, not all properties will necessarily exhibit a linear relationship with chain length and care must be taken in making assumptions about such trends. For many homologous series, increasing logKow leads to increasing fish toxicity whilst at the same time water solubility decreases. There is usually a point where the solubility is too low to be expressed. For example, in alpha-olefins there is an apparent cut-off point between the C8 and C10 chain length at which acute toxicity to fish is no longer observed. Similarly, a trend of increasing molecular weight may lead to decreasing systemic toxicity as absorption decreases. There may be a change of physical state of the category members as chain length increases.

Care should be taken when evaluating a category containing both branched chain chemicals and linear chain chemicals. Whilst there may be no influence of degree of branching on a trend for some endpoints (e.g. aquatic toxicity), significant differences could be expected for other endpoints (e.g. biodegradation). For these endpoints where differences in trend are seen, it may be helpful to divide the category into subcategories in order to provide a robust justification for the assessment.

Careful thought should be given to selecting the boundaries of a chain length category. The cut-off points described above may provide useful boundaries. The potential scope and size of a chain length category may be larger than that covered by a particular manufacturer or consortium. Where possible, well-characterised substances which are not necessarily HPV chemicals but which fit into the series should be included. There may be cases when testing the end members of a chain length category is not appropriate. For example if the existing data indicates that the toxicity cut-off occurs earlier in the series, it may not be necessary to test the end member for that endpoint.

QSARs can be used to help justify the category and fill data gaps. In general, substances at either end of a chain length category should have all endpoints fulfilled, preferably with test data. This permits interpolation of data to the other category members rather than extrapolation and increases confidence in the estimate. For example, in the category on ethylene glycols, a linear regression was used to predict acute

aquatic toxicity, indicating that toxicity decreases with increasing chain length, and further supporting the low toxicity of the category members concluded from available experimental data. For categories where there is more than one variable, such as variation in the length and degree of branching of the chains, more category members are likely to be required to bring confidence to the interpolations being made.

Other examples are oleochemical derivatives which can be grouped in such categories as fatty acids or alkyl sulfates. These categories may contain single-chain chemicals as well as mixtures containing chemicals of distinct chain lengths at varying amounts. The relative amounts of individual chain length molecules in mixtures are usually reflective of the chain length distribution in natural fats and oils from which they are derived. Since the category chemicals differ from each other only by the number of -CH_2-CH_2- units, these categories are very homogenous and exhibit a constant pattern in the changing of the potency of the properties across the category as described below.

6.2 Metabolic pathways

The underlying hypothesis for a metabolic series is a sequential metabolism of a parent chemical to downstream blood metabolites that are chemicals of interest. Hazard identification studies with the parent compound could then be used to identify the hazards associated with systemic blood levels of the downstream primary and secondary metabolites and once quantified, can be used in place of studies using direct exposure to primary and secondary metabolites themselves. In certain instances, the metabolism of the parent compound within barrier tissue (e.g. lung or gut tissue) occurs so rapidly that the initial primary metabolite is the predominant chemical found within the blood. Under these circumstances data from hazard identification studies conducted with that primary metabolite itself can be used to identify hazards for the parent compound. PBPK or PBPD models may help to define categories. The metabolic pathway approach is usually reserved to some toxicological endpoints. For physicochemical properties, environmental fate and ecotoxicity, information on the parent compound would need to be available. Examples of metabolic pathway categories which have been assessed within the OECD HPV Chemicals Programme are isobutyl isobutyrate (CAS No 97-85-8) or trimellitic anhydride (CAS No 552-30-7) (UNEP Chemicals, 2006).

The first technical issues faced when forming a metabolic series is to determine if the metabolism that is assumed to occur does occur independently of the requirements of the programme under which the chemical is assessed. This is necessary before moving any further in developing a metabolic category and preferentially should be determined in vivo. In certain instances, in vitro metabolic studies can be used to help identify metabolic pathways, but the definitive evidence should be conducted in whole animals. The primary and secondary metabolites should be detected either in the blood or tissue. Primary and secondary metabolites that cannot be readily determined in blood or tissue should not be candidates for a metabolic series approach without some limitation placed upon the use of the information.

The second technical issue pertains to the level of evidence required to describe the metabolic processes. Direct measurement of the parent chemical and primary and secondary metabolites in the blood in an in vivo exposure is the recommended standard. The level of evidence required to presume that there will be blood-borne levels of primary and secondary metabolites following exposure to parent chemical, will have to be determined on a case by case basis. Certain metabolic processes are ubiquitous and well understood and these can be presumed to occur without performing in vivo experiments in every instance. Other metabolic processes are not part of normal metabolism or require enzyme induction. These metabolic processes may not be well characterized and should not be assumed without specific in vivo evidence of blood levels of primary and secondary metabolites.

The third technical issue provides a limitation for the metabolic approach to forming categories. The metabolic category reasoning is only useful for identifying hazards related to systemic blood levels of the parent compound and/or primary and secondary metabolites. Other endpoints of hazard identification studies that are dependent upon site of contact effects (e.g. eye, skin, respiratory tract irritation, irritation to gastric mucosa) cannot be addressed using the metabolic category logic. These sites of contact effects are often due to the physic-chemical properties of the chemical in question and therefore may differ considerably between the parent compound and primary and secondary metabolites. In addition, tests that identify unique structural characteristics (e.g. skin or respiratory sensitisation) or are dependant upon physical chemical properties (e.g. volatility and LC_{50} values) should not be considered as part of metabolic category because these properties may not be similar amongst the various members of the metabolic series.

An additional limitation of the metabolic categories approach is that metabolism and toxicokinetics experiments have to be conducted with the parent compound. These types of studies are not requested in most review programmes and therefore would require a sponsor of the chemical to do additional work beyond what is normally considered necessary. However, it should be recognized that the savings involved (numbers of animals used, testing costs) could be considerable compared with generating data for each metabolic category member for each endpoint of systemic toxicity. For screening level assessments that are interested in identifying hazards related to systemic blood levels, it should not become necessary to provide definitive toxicokinetic evidence or develop a toxicokinetic model for acceptance of hazard identification studies as relevant for the primary and secondary metabolites.

An additional advantage of using the metabolic category toxicity data is that in certain instances, higher systemic blood levels of a chemical can be achieved from metabolic pathways than if the primary or secondary metabolite was administered directly. For example, if a material is corrosive or has limited volatility, higher blood levels may be found following the administration of the parent compound than if the primary or secondary metabolite was administered directly to the animal.

The following specific issues should be taken into account when developing a metabolic pathway category, according to the stepwise procedure described in Section 5.2.

a) Step 1: Provide definitive information on the metabolism of the parent chemical to the primary and secondary metabolite. This information should also include, preferably, time course data for either blood or tissue for both the parent chemical and the primary and secondary metabolites.

b) Step 2: The metabolism experiment should be examined to determine, if in fact, the primary and secondary metabolites are formed, if they achieve appreciable levels within the blood and/or tissues and determine basic toxicokinetic parameters for the parent material. For example, the $T_{1/2}$ for elimination for the parent chemical should be determined if possible. If the metabolism of the parent chemical to the primary metabolite is rapid and is thought to occur within barrier tissues, then it may be appropriate to use hazard identification studies from the primary metabolite to identify hazards associated with exposure to the parent chemical.

c) Step 3: If there are appropriate hazard identification studies that have been conducted with the parent chemical or primary or secondary metabolites for similar toxicity endpoints, then these studies should be examined to see if these materials have similar toxicity. If data are not available for the metabolic series in question and a study is to be designed and conducted, then the parent compound should be tested, so that blood levels of all category members will be present. The toxicokinetic and metabolic experiments that provide the basis for the metabolic category should have robust summaries prepared and be included in the dossier for the parent

chemical, primary and secondary metabolites. A table should be included detailing the relative blood levels of the parent chemical, primary and secondary metabolites.

d) Step 5: A quantitative analysis between exposures of the parent chemical and the primary and secondary metabolite is usually not necessary if the only objective is hazard identification. It is recognised that in certain cases quantitative differences play an important role in hazard identification (e.g. in the metabolism of C6 - C8 alkanes). For risk assessment purposes, a quantitative analysis may become necessary, e.g. additional toxicokinetic analysis (including preparing a model) may be appropriate.

The metabolic approach should not be used for environmental toxicity endpoints unless the metabolism of the parent compound to the primary or secondary metabolite can be demonstrated within the test species in question. Whereas it may be appropriate to extrapolate within mammals, it may not be appropriate to extrapolate between amphibia and fish or insects and other species due to the difference in the metabolic processes and enzymes present within those species.

On the other hand the same concept underlying the metabolic pathways can be used for environmental degradation processes. For example, for a substance which hydrolyses very rapidly in aquatic test systems (half-life < 1 hour), the aquatic toxicity endpoints can be covered by the test results with the degradation product(s) (OECD, 2000).

6.3 Chemical reaction products and multi-constituent substances

Categories can be developed for series of chemical reaction products or multi-constituent substances (MCS) that are related in some regular fashion. As with categories based on discrete chemicals, in a category containing reaction products or MCS some, but not all, of the individual substances may require testing.

A number of categories assessed under the OECD HPV Chemicals Programme provide useful case studies on dealing with multi-constituent substances. Further information is available at (http://cs3-hq.oecd.org/scripts/hpv/). For the Ethylene Glycols category, data from PEG 200, a mixture of chain lengths, was used to support the human health assessment. For the Linear alkylbenzene sulfonates category, aquatic toxicity data was available for both commercial products (mixtures) and pure C13 and C14 homologues. The pure homologues showed higher toxicity than the commercial mixtures but data for the pure homologues was not used to drive the recommendation of the assessment since they were not commercially supplied (Caley J et al., 2007). The Bicarbonate Special category focusing on ammonium bicarbonate, provided an interesting example of assessing a reaction mixture using data from pure components. The commercial material is a reaction mixture of sodium bicarbonate, sodium carbonate and ammonium bicarbonate. Aquatic toxicity data was available for the three components. Ammonium bicarbonate is the most toxic and the evaluation therefore focused on the quantity of ammonium ions released to water from dissolution of Bicarbonate Special and the impact of pH on the ammonium speciation and toxicity (Caley J et al., 2007). Effectively, the ammonium ion was used as a marker for aquatic toxicity (see also Section 6.5).

Another example is the reproductive toxicity of technical C7-C9 phthalate ester mixtures. In case of ortho phthalate esters, there was clear evidence that phthalates with a C4-C6 backbone (i.e. the length of the longest branch in the side chain) were reprotoxic, whereas phthalates with a backbone >C6 might not be. It was assumed therefore that phthalate ester mixtures which contained both lower and higher

homologues, then the reprotoxic capacity/potency of the mixture would depend on the amount of the lower homologues (backbone C4-C6) present in the mixture. In fact what was observed for some complex mixtures containing a high amount of the lower homologues was similar but fewer reprotoxic effects, at higher concentrations and with less severity than the lower homologues. Therefore, when assessing such mixtures, it would not be sufficient to determine just the predominant homologue or different homologues (side-chain, backbone lengths) in the mixture, but also the amount and properties of these different homologues (Fabjan et al., 2006).

The composition and physicochemical properties of substances are useful considerations to take into account when dealing with MCS.

6.4 Isomers

Isomers are chemicals that have identical molecular formulas but different molecular arrangements. Although there are several types of isomers, the two that typically will be considered are structural and geometric.

Structural isomers are molecules with differences in the arrangement of their atoms. Structural isomers can include:

a) chain isomers. For example hydrocarbon chains with identical or variable lengths and variable branching patterns (see also section 6.1).
b) positional isomers. For example hydrocarbon chains with a functional group that varies in position along the chain. An example is 1-butene and isobutene.
c) functional group isomers. These isomers also have identical molecular formulas, but contain different functional groups. Examples are 1-butanal and 2-butanone which both have the molecular formula $C_4H_{10}O$. Each of these isomers contains a carbonyl group (C=O), but are representative of two different chemical families: butanal is an aldehyde whereas butanone is a ketone. This type of structural isomers is less likely to be considered within a category because functional isomers can have very different chemical and biological properties. Functional isomers are not included within the scope of this guidance.

Stereoisomers are isomeric molecules whose atomic connectivity is the same but whose atomic arrangement in space is different. One type of stereoisomerism is geometrical (cis-trans) isomerism.

Geometric (or cis-trans) isomers can occur when a double bond or a ring is present. Bond rotation is restricted in these types of structures, so atoms can be permanently on the same (cis) or on opposite (trans) sides of the bond. For example, *cis*-2-butene and *trans*-2-butene each have carbon groups on either side of a double bond, which cannot rotate, so the carbon groups are arranged on either the same side of the molecule (*cis*) or opposite sides of the molecule (*trans*).

Enantiomers are two stereoisomers that are related to each other by a reflection: they are mirror images of each other. Every stereocentre in one has the opposite configuration in the other. Two compounds that are enantiomers of each other have the same physical properties, except for the direction in which they rotate polarized light and how they interact with different optical isomers of other compounds, and how they interact with enzymes. In nature, only one enantiomer of most chiral biological compounds, such as amino acids, is present. As a result, different enantiomers of a compound may have substantially different biological effects.

An example showing a profound difference in the effects of enantiomers is the drug thalidomide, The optical "R" isomer is an effective sedative whereas the optical "S"- isomer is a teratogen causing serious birth defects in children to mothers using the drug during pregnancy.

Stereoisomers can have similar or different chemical or toxicological properties. Even though they may behave identically in many chemical reactions, it is for example well known that the enzyme specificity in biological systems may be totally different, so caution is needed in case of such substances. An example of such specificity is certain carbohydrates, which may be metabolised or not depending on the orientation of functional groups. These are examples of diastereoisomers, which are defined as stereoisomers that are not enantiomers (i.e. they are not mirror images of each other). Diastereomers can have different physical properties and different reactivity.

There are two general principles for using estimation techniques as they apply to isomers:

a) Relatedness. The substance(s) with a data gap as well as substance(s) with data are similar such that their physicochemical, biological, and toxicological properties would be expected to behave in a predictably similar manner or logically progress across a defined range. This similar manner or logical progress should be demonstrated by the available experimental data. QSAR models and trend analysis can also be used in addition of experimental data to support the estimate.
b) Structural Similarity. The substance(s) with data gap possesses a small incremental structural difference from the reference substance(s) or the difference between the two would not be expected to affect the property sufficiently such that it could not be accurately predicted.This similar property should be demonstrated by the available experimental data. QSAR models and trend analysis can also be used in addition of experimental data to support the estimate.

There can be instances within a category of structural isomers when the estimate for an endpoint is not appropriate. An example is illustrated with two categories of isomers: the pentanes and hexanes. Although the pentanes may be broadly described as isomers, they actually represent three types of hydrocarbons, normal alkanes, branched alkanes, and cyclic alkanes. It is known that n-pentane, 2-methylbutane, 2,2-dimethylpentane, and cyclopentane exhibit distinct differences in potential biodegradability. n-Pentane and 2-methylbutane are readily biodegradable, whereas 2,2-dimethylpentane and cyclopentane are poorly biodegraded. Therefore, it is not possible to assess the biodegradability of the poorly biodegradable pentanes by using the results from the readily biodegradable pentanes, even though the pentane isomers could still be considered a category for other endpoints. In such a case, the potential biodegradability of the two groups of pentanes would each have to be characterised separately within the context of the category. Likewise, the peripheral neurotoxicity in humans associated with exposure to n-hexane has not been demonstrated to occur with exposure to other hexane isomers. Therefore, a discussion of this effect within a hexane isomer category would have to isolate n-hexane from the other isomers.

Based on the category of butenes and their mixtures, the following general principles were derived:

a) selected properties of isomers may be read-across to another isomer(s) or to an isomeric mixture within a category if the data are similar and/or if the structure of the isomer(s) without data is similar to the isomers with data.
b) extrapolating properties to isomeric mixtures should take into account mode of action, potential additivity and synergy, as well as purity profiles, and mixture composition.
c) for toxicological endpoints (e.g. LC_{50}, NOAEL), a range of toxicity or the lowest value in a range of toxicity may be used for read-across.

d) read-across from one isomer to another may not be straightforward. Metabolic data may be needed if existing knowledge of category members or related non category members suggests that differences may be expressed within a biological endpoint of interest.

6.5 Complex substances (UVCB)

Complex substances include a diverse range of materials which are defined as "substances of *Unknown or Variable composition, Complex reaction products or Biological material* (UVCB substances)". The range of different types of UVCB is very wide and the specific properties may be diverse, such that the applicability of a common approach needs justification. The following section highlights the key issues, however, it is recognised that in some sectors this approach has been more widely used than others and thus there needs to a cautious approach to defining categories and applying the following recommendations. There are many different types of complex substances, although generally they all have the following characteristics in common.

a) they contain numerous chemicals (typically closely related isomers and/or chemical classes with defined carbon number or distillation ranges), and cannot be represented by a simple chemical structure or defined by a specific molecular formula.
b) they are not intentional mixtures of chemicals.
c) many are of natural origin (e.g., crude oil, coal, plant extracts) and cannot be separated into their constituent chemical species.
d) the concept of "impurities" typically does not apply to complex substances.
e) they are produced according to a performance specification related to their physicochemical properties.

While CAS numbers are important for identifying substances, in the case of complex substances they do not represent a unique chemical and the specificity of the CAS number definition may vary (some CAS number definitions are rather narrow, some are very broad), e.g. CAS numbers for:

a) petroleum complex substances are based on a hierarchy of considerations including hydrocarbon type, carbon number range, distillation range and the last processing step,
b) coal derived complex substances are based on the applied production process and may include information on the distillation range and the chemical composition, and
c) NCS: natural complex substances (e.g., essential oils) are assigned CAS numbers based on their genus and species, in some cases part of plant, extraction method and other processing descriptors

Due to these numerous considerations, similar products sometimes have different CAS numbers. There are also historical and geographical reasons why similar complex substances may have been assigned different CAS numbers. Further, some CAS numbers have a broad definition that may fit different, but related complex substances that fall into different categories. These complexities lead to the use of physical properties and chemical descriptors (e.g. chain length, chemical class, size of aromatic ring systems) as being the preferred way to define categories of complex substances. In the case of NCS, this categorisation may also occur around the major chemical component(s) present, and might include marker chemicals for toxicity where it is clear that the behaviour of the UVCB is driven by those marker chemicals.

The approach used to define a category of complex substances may vary, although generally the approach will be related to how the category members are manufactured, defined and used.

6.5.1 General guidance on developing categories for complex substances

The key step is to define the category and identify category members. While initially this may seem repetitive, in fact the steps are different for complex substances. This is best explained by considering the "define analogue(s)" step, which for complex substances means identifying single component substances that represent the range of properties and the matrix being built up by the complex substances. The properties of these analogues are used, often with properties of the complex substances, to develop the data matrix and describe the physicochemical space.

The following elements are considered to be the main blocks to be used when putting together a category for complex substances.

(a) Composition - it is important to clearly characterise the complex substances to the extent it is measurable. In particular, it is necessary to identify which of the following attributes are key and must be specified:
- Cut off ranges
  - Range of chain length or predominant carbon number range or size of condensed ring systems
  - Distillation temperature range
  - Appropriate measures that allow characterisation of category members
- Known or generic composition and description
- Standard index – e.g. Colour Index number
- Chromatographic and other physical "fingerprints"
- Reference to standards
- Information on the production process (especially useful in categorising petroleum or coal derived products)
- For botanical NCS identification of the genus/species, origin should be considered
- If marker chemicals are appropriate, they should be clearly identified and if possible quantified for all category members.

(b) Properties of the components of a complex substance can be applied to the complex substance if the properties of the single components are similar, or fall within an expected range, depending on the endpoint.
- it is necessary to identify representative components of the complex substance to cover the carbon range and structure types of members of the complex substance.
- components with outlying properties need to be identified (e.g. specific toxicity of hexane compared to other aliphatic hydrocarbons, higher water solubility of aromatic hydrocarbons compared to aliphatic hydrocarbons).

(c) Data gap filling - Read-across/SAR and QSAR
It is possible to fill data gaps within a defined category either using read-across/SAR or establishing a QSAR, which is sometimes best described as a local QSAR. Where the composition of two, or more, complex substances is similar (within boundaries defined by the category description) qualitative properties can be established and data gaps filled. Quantitative read-across is more difficult in such circumstances, although it is possible to establish ranges. Where a valid QSAR is either available or can be established based on components of the complex substance, it can be possible to fill data gaps with either qualitative or quantitative information. When this is done justification for the approach and chosen data needs to be clearly described.

It is also very important to carefully consider the dose-response relationship for read-across/QSAR

versus the nature of the complex substances and the level of components of concern within the complex substances.

(d) Data gap filling – testing
Where it is necessary to identify representative complex substances for testing purposes, this should be done bearing in mind the key components of the category definition and the ranges thus defined.

6.5.2 Petroleum complex substances

Petroleum complex substances are generally defined by manufacturing and processing conditions, hydrocarbon chemistry (e.g., aliphatic hydrocarbons, aromatic hydrocarbons), physico-chemical properties such as boiling range or carbon-number range, and common use categories. An example of the grouping of petroleum complex substances, developed for the purposes of the Existing Substances Regulation and also used for classification and labelling purposes, is given in Comber M & Simpson B (2007). According to this approach, petroleum complex substances are grouped according to the process by which they are manufactured, on the assumption that substances within each group (or sub-group) have similar physicochemical properties and therefore similar intrinsic hazard properties. Within this approach, two substances and a class of chemicals (DMSO extractable PAHs) were used as markers for carcinogenicity, i.e. the presence of one of these substances at a specified level was used to indicate and classify for carcinogenicity. For other classification endpoints read-across between members of the categories has been used and more recently supported by QSAR.

The approach adopted for the petroleum complex substances has more general applicability to UVCBs and should be considered by other industries for which it may be applicable.

6.5.3 Hydrocarbon solvents

Hydrocarbon solvent categories are based on typical chemistry and carbon-number range. Common use can also contribute to the category definition. Under this approach, those hydrocarbon solvent substances with similar chemistry and carbon-number range are grouped within a category that is generally defined by the predominant constituents of the category members. This approach is practical and has the benefit of ensuring that similar commercial products are grouped together in the same category.

6.5.4 Coal derived complex substances

The principle described in 6.5.2 for petroleum derived complex substances also applies to coal derived complex substances. The longer geological history of coal compared to crude oil explains the higher degree of cross-linking of coal derived constituents. This results in a predominance of aromatic ring systems in coal derived complex substances. Longer alkyl chains do not appear. Processing of a coal derived feedstock separates according to volatility (size of condensed ring systems) and/or the extractability of acidic/ alkaline constituents. Formation of categories makes use of the applied processing techniques and of a similar spectrum of intrinsic properties for substances having a similar matrix of physicochemical properties.

6.5.5 Natural complex substances (NCS)

NCS are botanically-derived substances obtained by subjecting specific parts of the plant to a physical treatment such as extraction, distillation, expression, fractionation, purification, concentration or

to fermentation. Their compositions vary depending on the genus, species, the growing conditions and maturity of the crop used as a source, and the process used for its treatment.

NCS constitute a very specific subgroup of UVCBs (substances of unknown or variable composition, complex reaction products or biological materials) and include primarily essential oils and extracts obtained by various separation techniques.

Inclusion in a chemical group is possible based on the constituents of the NCS where the major components can be clearly identified as the same as known chemical substances. An example is provided in Salvito D (2007).

6.5.6 Use of toxic equivalency factors or toxic units approach for filling data gaps

The use of toxicity equivalency factors and the estimation of toxic units for mixtures of chemicals which contribute to a biological effect through a common toxicity pathway is a useful approach for filling data gaps in the assessment of chemical mixtures. The techniques are applied to mixtures of compounds in order to express the mixture's toxicity as a single value. The principle requirement is that the chemicals in the mixtures are active in a common toxicity pathway, and so this approach is strictly only applicable for chemical mixtures that have been formally grouped based on mechanistic considerations. Furthermore, toxicity data for the endpoint being assessed must be available for each component in the mixture.

Complex mixtures of PCBs (Clemens et al., 1994), furans (Parrott, 1992), dioxins (Safe, 1991; van der Weiden, 1992) and aromatic hydrocarbons (Walker, 1991; Zabel, 1995) have been assessed using toxicity equivalency factors based on Ah receptor binding and joint toxicity models amongst others. Joint toxicity models for calculating the toxic units generally use a strict addition model when a common toxicity pathway is a reasonable approximation. Although synergist effects are conceivable, they are only observed when chemicals in a mixture have different mechanisms, which should not be the case within a chemical category rigorously formed by the principles including toxic mechanistic considerations.

In the Toxic Equivalents (TEQ) approach, the most toxic compound is used as the reference compound. This compound does not necessarily have to be present in the mixture being assessed, but the components of the mixture must all act by the same single toxic pathway and be of the same compound type (structural/functional group similarity) as the reference. The components of the mixture are each assigned toxic equivalency factors (TEFs) such that their individual toxicity is expressed as a fraction of the toxicity of the reference compound (which is given a TEF of 1). This is achieved simply by dividing the effect value of the reference compound by the effect value of the particular component (equation 1).

$$\text{TEF (component A)} = \frac{\text{Reference effect value}}{\text{Component A effect Value}} \qquad \text{Equation 1}$$

The amount of each component in the mixture is then multiplied by its respective TEF and the values for each component are summed to give the overall toxic equivalency, relative to the reference compound (equation 2).

$$\text{TEQ} = \Sigma \text{ (concentration x TEF)} \qquad \text{Equation 2}$$

For example in the case of dioxin and furan mixtures, toxicity relative to 2,3,7,8-tetraCDD (2,3,7,8-tetrachloro-*p*-dioxin) was derived, based on mortality of rainbow trout fry following injection of the compounds to eggs. The following table lists TEFs derived from measured toxicity data for some of the compounds found in the literature (Safe, 1991, Walker, 1991, Zabel, 1995):

Dioxin/Furan	**Toxic Equivalency Factor**
2,3,7,8-tetraCDD	1 (reference compound)
1,2,3,7,8-pentaCDD	0.73
1,2,3,7,8,9-hexaCDD	0.1
1,2,3,6,7,8-hexaCDD	0.024

To illustrate the approach using a fictitious example based on these data:

Mixture A contains 20% 2,3,7,8-tetraCDD, 50% 1,2,3,7,8-pentaCDD, 10% 1,2,3,7,8,9-hexaCDD and 20% 1,2,3,6,7,8-hexaCDD.

Therefore, according to equation 1: (0.2 x 1) + (0.5 x 0.73) + (0.1 x 0.1) + (0.2 x 0.024) = 0.5798

So the toxic equivalency of Mixture A relative to the reference compound 2,3,7,8-tetraCDD is 0.5798, the fraction indicating a lower level of toxicity. In order to quote this fraction as an effect value (for example as an acute LC50 value) for Mixture A, the effect value of 2,3,7,8-tetraCDD is divided by 0.5798 giving a higher effect value (i.e. lower toxicity) for the mixture.

An adaptation of the method has been applied in the draft risk assessment of coal tar pitch (under the EU Existing Substances Regulation, CAS 65996-93-2 Pitch, coal tar, high-temp, EC, 2006c) in which the local concentration (C_{local}) for each component is divided by the component's PNEC, the summation of all expressing the risk characterisation ratio as opposed to toxicity (equation 3). A value greater than 1 indicated a risk.

$$\text{Sum RCR} = \Sigma \frac{C_{local}}{PNEC} \qquad \text{Equation 3}$$

In another adaptation of the method, the OECD HPV assessment of C6-22 Aliphatic Alcohols (Long Chain Alcohols, see http://cs3-hq.oecd.org/scripts/hpv/), measured acute fish toxicity data were not available for all of the alcohols present in these complex mixtures. Therefore (Q)SAR estimation was used to fill toxicity data gaps and so predict the toxicity of the complex mixtures.

In summary, toxic equivalency can be used for complex mixtures when there is a common mode of toxic action such that the effect is additive across the components of the mixture: there is no synergism. In addition, measured toxicity data should be available for each individual component of the mixture. Differences in test protocol for each data point can have a marked effect on the derived TEFs (and so TEQ), therefore if this approach is followed then it is necessary to present all available data and justify the use of the approach. This includes discussion of the shared toxic mechanism of the components in the mixture, choice of data for deriving the TEFs, discussion of the purity of the mixture/presence of impurities and their effects, and any deviations from the method.

6.6 Metals, metal compounds and other inorganic compounds.

The concept of chemical categories has traditionally been widely used for hazard assessment for certain endpoints and risk assessment of inorganic substances. The approaches have generally been based on the occurrence of a common metal ion or anion and the use of read-across to fill the data gaps.

For example, the chemical category approach based on the metal ion has been extensively used for the classification and labelling of metal compounds in the EU[13]. Other category entries are based on certain anions of concern such as oxalates and thiocyanates. For these EU classifications the category approach has often been applied to certain endpoints of particular concern for the compounds under consideration, and has not necessarily been applied to all endpoints of each individual compound in the category of substances. A category approach has also been used during the categorisation of existing chemicals on Canada's domestic substances list (Environment Canada, 2003).

This approach has also been used for estimating the potency of the effects as well as for their identification. NOAEL(s), NOEC(s) and comparable quantitative estimates have been read-across from data obtained from water-soluble compounds to other water-soluble compounds, including, in the absence of specific data, to compounds of substantially lower water-solubility. One example is the EU risk assessments on nickel (Tsakovska I & Worth A, 2007).

The application of these concepts has been useful[14]

- to evaluate hazards for substances for which data are limited rather than relying exclusively on conducting tests.
- to evaluate hazards for a range of compounds regarded as "difficult" substances, as they can present technical difficulties when carrying out standard test protocols […].
- to evaluate hazards for a number of metal compounds, for which animal models do not always reliably predict effects on humans. Where the hazard has been identified on the basis of human data the use of read-across provides a method to avoid these difficulties.

The guidance below is based largely on the practice of the EU Technical Committee on Classification and Labelling, the EU Technical Committee on New and Existing Substances and experience gained in other fora (see also Hart J, 2007 and Schoeters I & Verougstraete V, 2007 for examples). This guidance is intended to supplement the general guidance in the previous chapters with issues specific to metals and inorganic compounds.

6.6.1 Assumptions underlying the grouping of metal compounds

There are a number of assumptions underlying any grouping of metal compounds for estimating their biological properties.

The hypothesis is that properties are likely to be similar or follow a similar pattern as a result of the presence of a common metal ion (or ion complex including a hydrated metal ion). This is a reasonable

[13] The EU terminology for this type of entry is a "group entry" rather than a category.

[14] The approach of grouping metals and metal compounds in risk assessments has also been applied because it allows addressing together all compounds which potentially lead to exposure to the same metal moiety.

assumption for the majority of inorganic compounds and some organic compounds (e.g. metal salts of some organic acids). However, it is the bioavailability of the metal ion (or a redox form of this ion) at target sites that in most cases determines the occurrence and severity of the effects to be assessed for the read-across of metal substances. Supporting information to assess the bioavailability of the metal ion at the target site can include information on a number of different factors (e.g. physicochemical properties such as water solubility, degree of dissociation of the metal –containing compound, particle size and structure[11], *in vitro* solubility, *in vivo* data on systemic effects, toxicokinetics).

6.6.2 Basis for the development of categories or read-across approach of metal compounds

Hazard data is available for some primary metals and some key (high production volume) inorganic compounds. However, for a wide range of inorganic and organic compounds of the same metal, data is usually very limited. Data availability will play an important role in the selection of source chemicals.

As metals occur in a wide and heterogeneous range of substances, including inorganic metal compounds, organic metal salts, organometallic compounds, metals, metal-metal compounds (i.e. compounds containing more than one type of metal), alloys and complex substances, care is needed in order to select those metal compounds for which a category approach is relevant from those where read-across is not applicable.

The following points could alter the assumption of commonality and should be considered:

- Chemical speciation and valency

  When selecting the appropriate source substance, the valence state and its influence on the assumption of commonality should be checked. For some metals (predominantly transition elements), the chemical speciation and in particular the different valencies may result in differences in mechanism of action and a variation in toxicological properties. For example, differences in hazards are seen with Cr^{3+} and Cr^{6+} compounds. In some cases, species may be interconvertible, in other cases there is little interconversion between the species.

- Organometallic compounds

  Organometallic compounds will generally have a different mode of action since the metal ion is not likely to be present in the same form as for inorganic compounds. In such cases, read-across between inorganic and organometallic compounds is not recommended, although read-across may well be appropriate between different organometallic compounds. On the other hand, especially for environmental risk assessment, if an organometallic compound degrades rapidly to its inorganic metal moiety, it can be assessed together with the inorganic metal moiety

- Metals

  Particular difficulties have been seen in evaluating the properties of metals on the basis of data for metal compounds. In some cases, read-across of properties from the metal compounds to the metal itself (metallic, zero-valent form) has been agreed (e.g. cadmium oxide to cadmium metal, EC 2006a), whilst for others it has not (e.g. soluble nickel salts to nickel metal, EC 2006b). These need to be evaluated on a case-by-case basis.

- Metal containing UVCBs

Some metal containing UVCB compounds may not be appropriate for consideration in a category approach, as their effects will not be expected to be adequately described by their metal content. These include compounds such as asphalt, frits and drosses. In cases where read-across is not considered appropriate, clear arguments should be put forward as to why the known hazard profile of the metal is not expected to be relevant (for example very low bioavailability).

- Crystalline structure

The crystalline structure of insoluble metal compounds could influence the hazard profile. If there is reason to believe that the crystalline structure influences significantly the effects of the compound to be assessed, this must be taken into account in the evaluation. An example is silica of which the crystalline and non-crystalline forms have a different hazard profile (see category for syntheticamorphous silicas assessed within the OECD HPV Chemicals Programme; Silicon dioxide [CAS Nos 7631-86-9, 112945-52-5, 112926-00-8] Silicic acid, aluminum sodium salt [CAS No 1344-00-9] Silicic acid, calcium salt [CAS No 1344-95-2]).

6.6.3 Preliminary evaluation of the category and read-across

The water solubility of the metal compounds is often used as the starting point for establishing a category, as this provides a first indication of the availability of the metal ion in the different compartments of interest. For example, for inorganic nickel a number of sub- categories have been suggested, reflecting different ranges of aqueous solubility (Hart, J 2007).

The most simplistic approach to hazard evaluation is to assume that the specific metal-containing compound to be evaluated shows the same hazards as the most water-soluble compounds. This is a conservative approach, since systemic metal ion availability will normally be reduced with decreasing water-solubility and consequently reduced bioavailability.

This simplistic approach can be refined for categories containing many substances by building subcategories based on water solubility, when data is available on trends with water solubility. For example, mixed oxides with limited water solubility can be evaluated by comparison with the hazard profile for the metal oxides (where this is known) rather than for the soluble salts.

This difference in trend is clearly recognised in evaluating the environmental hazards of metals and metal compounds, where the relevant hazards can be evaluated using a transformation/dissolution protocol (OECD 2001).

Information from other endpoints could further support the systemic bioavailability assumptions. For example, the LD_{50} values for the semi-soluble nickel compounds was used to demonstrate systemic uptake to justify classification for reproductive toxicity for these compounds, but not for the less soluble oxides and sulfides (Hart J, 2007). For endpoints where a threshold occurs, estimates of the systemic bioavailability (i.e. toxicokinetics) of the metal ion can be ascertained for representative members of each category in order to ascertain whether the bioavailability exceeds the threshold for the compounds.

In addition to water solubility, phagocytosis, bioaccessibility in synthetic biological fluids, and organ deposition and clearance rates are relevant parameters to be considered (Schoeters I & Verougstraete V, 2007).

Where toxicokinetic data is available, this should be used as this provides relevant information on whether the source and target chemicals in question behave similarly as expected from read-across or whether there are biologically differences that would bring into question the validity of the category hypothesis.

Other factors may also need to be taken into account.

- Counter ions and other metal ions:

  The assumption that the metal ion is responsible for the common property or effect implies that the toxicity of the counter ion or of other metals present in the compound will be largely irrelevant in producing the effects to be assessed. This assumption could be affected by interactions between the metal ion and other parts of the substance e.g. the counter ion. It is noted that in certain cases the effect of the counter ion in acute toxicity studies exert another effect than in repeated dose studies using lower dose levels. This could obscure the role of the metal ion in either the acute or repeated dose studies. The influence of the counter ion should be checked for each endpoint. If there is reason to believe that the counter-ion (such as cyanates, oxalates) or other metal ions present in the compound influence significantly the effects of the compound to be assessed and alter the assumption of commonality, this must be taken into account in the evaluation. One option may be to use the additive approach described in the Foreword to Annex I, Directive 67/548/EEC, in the guidance to Note A. (See also 6.6.5 below).

- Crystalline structure:

  The crystalline structure of insoluble metal compounds could influence the hazard profile. If there is reason to believe that the crystalline structure influences significantly the bioavailability and so the effects of the compound to be assessed, this must be taken into account in the evaluation. An example is the low bioavailability of spinels and rutiles.

- Particle size information:

  Particle size information of the substance influences the deposition behaviour in the respiratory tract and potential toxic effects. Based on particle size distribution data, trends in deposition and potency of effects can be assessed for locally acting substances.

If there is evidence that the crystalline structure and particle size influence significantly the bioavailability and so the severity of the effects of the compound to be assessed, this must be taken into account in a weight of evidence approach considering all available information (e.g. toxicokinetics ...)

6.6.4 Considerations of the need for further refinement

As described previously, a preliminary assessment of the analogue approach or category approach should be carried out to determine whether the rationale is supported and whether the approach is sufficiently robust for the assessment purpose. If these criteria are satisfied for a particular endpoint, the data gaps can be filled according to the guidance in Chapter 3.

If these criteria are not satisfied (there is uncertainty or contradictory information), the assessor should consider what additional information may be required. Additional data could include demonstrating

a difference in bioavailability/bio accessibility between the substances in a proposed analogue or category approach.

The following options could be considered:

- *In vitro* data:

  In vitro information may be obtained by determining relative solubilities in physiological media (e.g. synthetic gastric juice, synthetic sweat) or by the use of the transformation/dissolution protocol (OECD 2001) for the endpoints of sparingly soluble metal compounds related to the aquatic environment.

  The solubility in alveolar liquids, lysosomal liquid, mucous liquids may provide more relevant information than simple water solubility for argumentation of the extent of availability of the soluble fraction of material during its dwelling time in various regions of the respiratory tract. To test whether slightly soluble, particulate metal compounds are taken up into mammalian cells and release metal ions intracellularly as free metal ions or bound to cellular macromolecules and whether the metal ions reach the cell nuclei, tests in vitro can be carried out using phagocytosing mammalian cells in culture.

- *In vivo* data:

  In some cases, *in vivo* testing may be considered, especially for endpoints where there is uncertainty about the role of the counter-ion. In planning the testing, a starting point for the studies should be confirmation of the effects expected on the basis of a read-across. As an example, if read-across would indicate the skin irritation is expected, an initial test could be carried out *in vitro* to confirm this effect before *in vivo* testing is considered.

- Toxicokinetic data:

  Animal model systems (using rats and mini-pigs) have been successfully used to characterise the speciation-dependent bioavailability differential for metals such as lead, arsenic and cadmium (US Environmental Protection Agency, 2004). Alternative strategies using rare stable isotopes of metals such as lead and zinc have been successfully used for the ascertainment of bioavailability of these metals in humans and animals. These types of studies are not requested in most review programmes and therefore would require an assessor to do additional work beyond what is normally considered necessary. However, where such information is not available, information could be collected for representative members of the category.

6.6.5 General guidance for other compounds

Similar considerations are expected to apply to salts in which the anion is associated with the toxic effects (e.g. cyanides, oxalates, thiocyanates). For categories that cover reactive chemicals, the reaction/degradation products must be of a similar nature for each member of the category to be plausible (Caley J et al., 2007). One example is the Methanolates category assessed under the OECD HPV

Chemicals Programme (http://cs3-hq.oecd.org/scripts/hpv). This consists of 17 potassium and sodium methanolate and both react rapidly in water to form the corresponding hydroxide.

When comparing acids and their salts, differences arising from pH effects should be considered (Caley J et al., 2007). For example, skin and eye irritation are likely to be different for an acid compared with its salt. This is illustrated by the Phosphonic Acid Compound (Groups 1, 2, 3) categories assessed under the OECD HPV Chemicals Programme (http://cs3-hq.oecd.org/scripts/hpv). For these categories, dermal and irritation studies are considered separately for the acid and salts.

For the Gluconates category assessed under the OECD HPV Chemicals Programme (http://cs3-hq.oecd.org/scripts/hpv), it was found that for categories including ionisable compounds, the effect of the counter-ion needs to be considered (Caley J et al., 2007). It is possible that the counter-ion(s) may pose hazards of greater concern than the common cation or anion on which the category is based (e.g. metal counter-ions that are inherently hazardous on their own).

Under such circumstances, it may be of limited utility to group and assess substances by the component which is expected to have the least effect. In other cases, it may be concluded that effects of the counter-ion are insignificant and therefore need not be taken into account in the assessment.

CHAPTER 7. REPORTING FORMATS FOR ANALOGUE AND CATEGORY EVALUATIONS

This chapter provides reporting formats for analogue and chemical category approaches. The documentation of an analogue category approach is an integral part of the assessment report and this chapter provides guidance on how to report the or chemical category approach in e.g. Chapter 1 of a SIDS Initial Assessment Report or Chemical Safety Report. An example is given in Appendix 1.

For chemical categories the assessment report should address all members of the chemical category and be accompanied for each member of the category by the dossiers containing robust study summaries of the key studies for all relevant endpoints (physical chemical properties, environmental fate and pathways, ecotoxicity, toxicity).

Experience in the OECD HPV Chemical Programme has shown that for a simple analogue approach (read-across), it can be more practical to perform separate assessment reports for the source and target chemicals. In this case, the guidance below is relevant for the target chemical only, provided that the assessment(s) and dossier(s) of the source chemical(s) are referenced. In case no assessment is performed for the source chemical(s), the assessment report and dossier of the target chemical should contain all the relevant information, including robust study summaries from studies performed with the source chemical(s).

Furthermore, when developing an analogue or chemical category approach with IUCLID 5 or any other similar software having implemented the OECD harmonised templates (OECD 2006b), dedicated fields are provided in the software where users can insert or append the documentation elaborated with the present formats. Specific guidance on how IUCLID 5 can be used to construct and document an analogue read-across or chemical category can be found in the IUCLID Manual (EC, 2007).

7.1 Reporting Format for analogue read-across

1) Hypothesis for the analogue approach

Describe the molecular structure a chemical must have to be suitable as a source chemical. All functional groups need to be identified. Provide the hypothesis for why the read-across can be performed. If there is a mechanistic reasoning to the read-across, describe the foreseen mode of action for source and target chemicals and if relevant describe the influence of the mode of administration (oral, dermal, inhalation).

List the endpoints for which the analogue approach is applied.

2) Source chemical(s)

Describe the source chemical(s) as comprehensively as possible. Provide CAS numbers, names and chemical structures of the source chemical(s).

3) Purity / Impurities

Provide purity/impurity profiles for the target and source chemicals, including the likely impact on the relevant endpoints. It should be discussed which influence these impurities are thought to have on physico-chemical parameters, fate and (eco)toxicology, and hence on the read-across.

4) Analogue approach justification

Based on available experimental data, including basic physicochemical properties, summarise how these results verify that the read-across is justified. The data should also show that functional groups not common to source and target chemicals do not affect the anticipated toxicity. The available experimental results in the data matrix reported under 5) below should support the justification for the read-across.

More detailed discussion of available test results for individual endpoints (i.e. discussion of the selection of key studies, variability of experimental results between source and target chemicals etc.) should be provided in the corresponding sections of the assessment report (e.g. chapters 2-4 of the SIDS Initial Assessment Report or chapters 4-7 of the Chemical Safety Report).

5) Data matrix

Provide a matrix of data (endpoints vs. target and source chemicals) (see Figure 5).

In each cell in the Data Matrix, the study result type should be indicated in the first line, e.g.:

- experimental result
- experimental study planned
- read-across from supporting substance (structural analogue or surrogate)
- (Q)SAR

If experimental results are available, the key study results should be shown in the Data Matrix.

Figure 5. Data Matrix, Analogue Approach

CAS #				
CHEMICAL NAME	[Target chemical]	[Source Chemical 1]	[…]	[Source Chemical n]
PHYSICAL-CHEMICAL DATA				
Melting Point				
Boiling Point				
Density				
Vapour Pressure				

Partition Coefficient (log Kow)				
Water Solubility				
…				
ENVIRONMENTAL FATE and PATHWAY				
Photodegradation				
Stability in Water				
Transport and Distribution				
Aerobic Biodegradation				
…				
ENVIRONMENTAL TOXICITY				
Acute Toxicity to Fish				
Acute Toxicity to Aquatic Invertebrates				
Toxicity to Aquatic Plants				
…				
MAMMALIAN TOXICITY				
Acute Oral				
Acute Inhalation				
Acute Dermal				
Repeated Dose				
Genetic Toxicity *in vitro* . Gene mutation . Chromosomal aberration				
Genetic Toxicity *in vivo*				
Reproductive Toxicity . Fertility . Developmental Toxicity				
…				

More detailed discussion of how data gaps are filled for individual endpoints should be provided in the corresponding sections of the assessment report (e.g. SIDS Initial Assessment Report or Chemical Safety Report).

7.2 Reporting Format for chemical categories

1) Category definition and its members

1.1) Category Definition

a) Category Hypothesis

Describe the molecular structure a chemical must have to be included in the category. Provide a brief hypothesis for why the category was formed: the hypothetical relational features of the category i.e. the chemical similarities (analogies), purported mechanisms and trends in properties and/or activities that are thought to collectively generate an association between the members. All functional groups of the category members need to be identified. If there is a mechanistic reasoning to the category, describe the foreseen mode of action for each category member and if relevant describe the influence of the mode of administration (oral, dermal, inhalation).

b) Applicability domain (AD) of the category

Describe the set of inclusion and/or exclusion rules that identify the ranges of values within which reliable estimations can be made for category members. Clearly indicate the borders of the category and for which chemicals the category does not hold. For example, the range of logKow values or carbon chain lengths over which the category is applicable. The justification for the inclusion and/or exclusion rules should be reported under section "2) Category justification" below.

c) List of endpoints covered

List the endpoints for which the category approach is applied. Also indicate if for some endpoints the category approach can only be applied to a subset of the members of the category (subcategories).

1.2) Category Members

Describe all category members as comprehensively as possible. Provide CAS numbers, names and chemical structures of all category members.

1.3) Purity / Impurities

Provide purity/impurity profiles for each member of the category, including their likely impact on the category endpoints. It should be discussed which influence these impurities are thought to have on physico-chemical parameters, fate and (eco)toxicology, and hence on the read-across.

2) Category justification

Based on available experimental data (including appropriate physicochemical data and additional test results generated for the assessment of this category) summarise how these results verify that the category is robust. This should include an indication of the trend(s) for each endpoint. The data should also show that functional groups not common to all the (sub)category members do not affect the anticipated toxicity. The available experimental results in the data matrix reported under 3) below should support the justification for the read-across.

More detailed discussion of available test results for individual endpoints (i.e. discussion of the selection of key studies, variability of experimental results between different members of the category etc.) should be provided in the corresponding sections of the assessment report (e.g. chapters 2-4 of the SIDS Initial Assessment Report or chapters 4-7 of the Chemical Safety Report).

3) Data matrix

Provide a matrix of data (category endpoints vs. members). It should be constructed with the category members arranged in a suitable order (e.g. according to molecular weight) (see figure 6). For example, the ordering of the members should reflect a trend or progression within the category.

In each cell in the Data Matrix, the study result type should be indicated in the first line, e.g.:

- experimental result
- experimental study planned
- read-across from supporting substance (structural analogue or surrogate)
- trend analysis[15]
- (Q)SAR

If experimental results are available, the key study results should be shown in the Data Matrix.

Figure 6: Data Matrix, Chemical Category

CAS #					
CHEMICAL NAME	[Category member 1]	[Category member 2]	[Category member 3]	[…]	[Category member n]
PHYSICAL-CHEMICAL DATA					
Melting Point					
Boiling Point					
Density					
Vapour Pressure					
Partition Coefficient (log Kow)					
Water Solubility					
…					
ENVIRONMENTAL FATE and PATHWAY					
Photodegradation					
Stability in Water					
Transport and Distribution					
Aerobic Biodegradation					
ENVIRONMENTAL TOXICITY					

[15] There are slight differences between the terminology used in the OECD Harmonised templates and hence there might be slight differences in a category matrix automatically generated with software using the OECD Harmonised Templates and the present guidance document. For example there is no item "trend-analysis" in the picklist for the data element "study result type". Instead the item "read-across based on grouping of substances (category approach)"could be used.

Acute Toxicity to Fish					
Acute Toxicity to Aquatic Invertebrates					
Toxicity to Aquatic Plants					
…					
MAMMALIAN TOXICITY					
Acute Oral					
Acute Inhalation					
Acute Dermal					
Repeated Dose					
Genetic Toxicity *in vitro* . Gene mutation . Chromosomal aberration					
Genetic Toxicity *in vivo*					
Reproductive Toxicity . Fertility . Developmental Toxicity					
…					

Footnote to Figure 6: For data-rich substances, the matrix could become very large, and could therefore be broken down into groups of endpoints.

More detailed discussion of how data gaps are filled for individual endpoints and individual category members (e.g. interpolation, extrapolation, (Q)SAR) as well as the rationales for the chosen method of filling the data gaps should be provided in the corresponding sections of the assessment report (e.g. chapters 2-4 of the SIDS Initial Assessment Report or chapters 4-7 of the Chemical Safety Report).

For UVCBs it may not be feasible to establish a full data matrix, especially where the number of substances in the category is very large. In such circumstances a single data set or template that applies to all members of the category of UVCBs in exactly the same way will be developed. The template will include a clear indication of which members of the category experimental or calculated data exist, and hence maintain complete transparency.

APPENDIX 1. CASE STUDY USING PHOSPHONIC ACID COMPOUNDS AND ALKALI METAL SALTS

This case study is based on an assessment performed within the OECD HPV Chemicals Programme. For further details see http://cs3-hq.oecd.org/scripts/hpv/ (category phosphonic acids).

1) Category definition and its members

1.1) Category Definition

a) Category Hypothesis

Describe the molecular structure a chemical must have to be included in the category. Provide a brief hypothesis for why the category was formed: the hypothetical relational features of the category i.e. the chemical similarities (analogies), purported mechanisms and trends in properties and/or activities that are thought to collectively generate an association between the members.

This category covers 1-Hydroxy-1,1-ethane-diphosphonic acid (HEDP) and various sodium and potassium salts of that acid. The different salts are prepared by neutralising the acid to a specific pH. All category members are based on the HEDP structure, which can be de-protonated up to 5 times.

The category hypothesis is that all the members are various ionised forms of the acid 2809-21-4. The main assumption is that sodium and potassium are not significant in respect of all the properties under consideration. In dilute aqueous conditions of defined pH a salt will behave no differently to the parent acid, at identical concentration of the particular speciated form present and will be fully dissociated. Hence some properties (measured or expressed in aqueous media, e.g. ecotoxicity) for a salt can be directly read across (with suitable mass correction) to the parent acid and vice versa. Where dermal or irritation studies are available the acid and salts are considered separately.

The properties of HEDP and its salts are profoundly directed by their ionisation behaviour and complexation of metal ions.

b) Applicability domain (AD) of the category

Describe the set of inclusion and/or exclusion rules that identify the ranges of values within which reliable estimations can be made for category members. For example, the range of LogKow or carbon chain lengths over which the category is applicable.

The category applies to HEDP and all of its possible sodium and potassium salts.

c) List of endpoints covered

List the endpoints for which the category approach is applied. Also indicate if for some endpoints the category approach can only be applied to a subset of the members of the category (subcategories).

The category approach was applied to the following endpoints:

- Dissociation constant and metal complexation
- Octanol-water partition coefficient
- Adsorption
- Biodegradation
- Stability in water
- Bioaccumulation
- Ecotoxicity tests
- Mammalian toxicity (other than dermal administration)
- Genotoxicity

The category approach was not applied to skin irritation, eye irritation and dermal toxicity since the acid is much more corrosive than its salts.

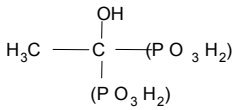
1.2) Category Members

Describe all category members as comprehensively as possible. Provide CAS numbers, names and chemical structures of all category members.

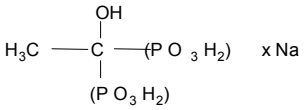
Table 1.1 Category Members

Substance	CAS
1-Hydroxy-1,1-ethane-diphosphonic acid	2809-21-4
1-Hydroxy-1,1-ethane-diphosphonic acid, xNa Salt	29329-71-3
1-Hydroxy-1,1-ethane-diphosphonic acid, Na Salt	17721-68-5
1-Hydroxy-1,1-ethane-diphosphonic acid, 2Na Salt	7414-83-7
1-Hydroxy-1,1-ethane-diphosphonic acid, 3Na Salt	2666-14-0
1-Hydroxy-1,1-ethane-diphosphonic acid, 4Na Salt	3794-83-0
1-Hydroxy-1,1-ethane-diphosphonic acid, 5Na Salt	13710-39-9
1-Hydroxy-1,1-ethane-diphosphonic acid, xK Salt	67953-76-8

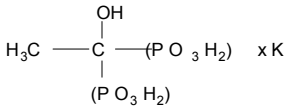
Substance	CAS
1-Hydroxy-1,1-ethane-diphosphonic acid, K Salt	17721-72-1
1-Hydroxy-1,1-ethane-diphosphonic acid, 2K Salt	21089-06-5
1-Hydroxy-1,1-ethane-diphosphonic acid, 3K Salt	60376-08-1
1-Hydroxy-1,1-ethane-diphosphonic acid, 4K Salt	14860-53-8
1-Hydroxy-1,1-ethane-diphosphonic acid, 5K Salt	87977-58-0



1-Hydroxy-1,1-ethane-diphosphonic acid
CAS # 2809-21-4



1-Hydroxy-1,1-ethane-diphosphonic acid, xNa Salt
CAS # 29329-71-3



1-Hydroxy-1,1-ethane-diphosphonic acid, xK Salt
CAS # 67953-76-8

1.3) Purity / Impurities

Provide purity/impurity profiles for each member of the category including their likely impact on the category endpoints.

Since the salts are prepared from the acid, the impurity profile for HEDP acid given in table 1.2 below is also typical of the salts in this Category, although acidic impurities would also be present as salts. Exact proportions vary slightly between manufacturers and precise values are not given, to protect commercial interests. All are typical for marketed substance. In addition to those impurities listed in Table 1.2, HEDP contains up to 4% of two phosphonic acid components, not unrelated to the main component. Exact details are commercially confidential.

Table 1.2 Impurity profile for HEDP

CAS-No	EC-No	EINECS-Name	Mol. Formula	Contents % w/w
64-19-7	200-580-7	Acetic acid	$C_2H_4O_2$	< 1
7647-01-0	231-595-7	Hydrogen chloride	HCl	< .1
13598-36-2	237-066-7	Phosphonic acid	H_3PO_3	< 4
7664-38-2	231-633-2	Orthophosphoric acid	H_3PO_4	< 2

2) Category justification

Based on available experimental data, including appropriate physicochemical data (including additional test results generated for the assessment of this category), summarise how these results verify that the category is robust, including an indication of the trend(s) for endpoint. More detailed discussion of available test results for individual endpoints should be provided in the corresponding sections of the assessment report (e.g. SIDS Initial Assessment Report or Chemical Safety Report).

HEDP and its salts all have high water solubility, low LogKow, and low vapour pressures. Their behaviour in water and biological systems is dominated by their ionisation and complexation of metal ions. Measured data was available for environmental endpoints for HEDP and its 2Na salt and for health endpoints for HEDP, its 2Na salt and 4Na salt. Thus, data is read-across to the remaining Na salts and to all potassium salts.

Data for HEDP and the 2Na salt showed low acute toxicity to fish, this result was read across to the remaining salts. Data for HEDP and the 2Na salt showed low acute toxicity to Daphnia, which was read across to the other category members. However, the available data indicated that the 2Na salt has a much higher chronic toxicity to Daphnia than HEDP. This result is not consistent with the general pattern of toxicity and therefore a repeat test was requested on the 2Na salt (result not yet available). If the test confirms the chronic toxicity of the 2Na salt, the category may be called into question for aquatic toxicity endpoints. Data for the toxicity of HEDP and its 2Na salt to algae shows toxicity, but evidence shows that these effects are a consequence of complexation of essential nutrients and not of true toxicity. This conclusion applies to the whole category.

3) Data matrix

Provide a matrix of data (category endpoints vs members). It should be constructed with the category members arranged in a suitable order (e.g. according to molecular weight). For example, the ordering of the members should reflect a trend or progression within the category. The cells of the matrix should indicate whether data are available or unavailable and the available key study results should be shown. For cells for which no reliable experimental results are available, indicate how the data gap is filled (read-across, trend analysis, (Q)SAR) and indicate the result.

More detailed discussion of how data gaps are filled for individual endpoints and individual category members should be provided in the corresponding sections of the assessment report (e.g. SIDS Initial Assessment Report or Chemical Safety Report).

Physicochemical properties and environmental fate - data which are read-across are highlighted in yellow

Substance	CAS	Water solubility	Log Kow	Vapour pressure	Melting point	pKa	Vapour pressure	Koc	biodegradability
1-Hydroxy-1,1-ethane-diphosphonic acid	2809-21-4	690 g/l: 60% w/w produced commercially	-3.52	1.24 x 10^{-9} Pa (estimated)	198-199° C; decomposes around 228° C	Four pKa values of HEDP (at 0.1 M ionic strength potassium nitrate): 1.6, 2.7, 6.9, 11.0.	1.24 x 10^{-9} Pa (estimated)	16610	Not readily biodegradable (NRB) (measured)
1-Hydroxy-1,1-ethane-diphosphonic acid, xNa Salt	29329-71-3	'high'	'low'	'low'	-	-	'low'	'high'	NRB – read across
1-Hydroxy-1,1-ethane-diphosphonic acid, Na Salt	17721-68-5	465 g/kg solution	'low'	'low'	-	-	'low'	'high'	NRB – read across
1-Hydroxy-1,1-ethane-diphosphonic acid, 2Na Salt	7414-83-7	278 g/kg solution	'low'	'low'	-	-	'low'	'high'	Not readily biodegradable (measured)
1-Hydroxy-1,1-ethane-diphosphonic acid, 3Na Salt	2666-14-0	123 g/kg solution	'low'	'low'	-	-	'low'	'high'	NRB – read across
1-Hydroxy-1,1-ethane-diphosphonic acid, 4Na Salt	3794-83-0	513 g/kg solution	'low'	'low'	-	-	'low'	'high'	NRB – read across
1-Hydroxy-1,1-ethane-diphosphonic acid, 5Na Salt	13710-39-9	'high'	'low'	'low'	-	-	'low'	'high'	NRB – read across
1-Hydroxy-1,1-ethane-diphosphonic acid, xK Salt	67953-76-8	'high'	'low'	'low'	-	-	'low'	'high'	NRB – read across
1-Hydroxy-1,1-ethane-diphosphonic acid, K Salt	17721-72-1	'high'	'low'	'low'	-	-	'low'	'high'	NRB – read across
1-Hydroxy-1,1-ethane-diphosphonic acid, 2K Salt	21089-06-5	'high'	'low'	'low'	-	-	'low'	'high'	NRB – read across

Substance	CAS	Water solubility	Log Kow	Vapour pressure	Melting point	pKa	Vapour pressure	Koc	biodegradability
1-Hydroxy-1,1-ethane-diphosphonic acid, 3K Salt	60376-08-1	'high'	'low'	'low'	-	-	'low'	'high'	NRB – read across
1-Hydroxy-1,1-ethane-diphosphonic acid, 4K Salt	14860-53-8	'high'	'low'	'low'	-	-	'low'	'high'	NRB – read across
1-Hydroxy-1,1-ethane-diphosphonic acid, 5K Salt	87977-58-0	'high'	'low'	'low'	-	-	'low'	'high'	NRB – read across

Ecotoxicity endpoints – data which are read-across are highlighted in yellow

Substance	CAS	Fish acute toxicity 96h LC50 mg/l	Daphnia acute toxicity 48h EC50 mg/l	Daphnia chronic toxicity 22d NOEC mg/l	Algal toxicity 96h EC50	Algal toxicity NOEC mg/l	toxicity to microorganisms 30-min EC0 mg/l
1-Hydroxy-1,1-ethane-diphosphonic acid	2809-21-4	200	167	6.75 (28-day)	3	13 (14d)	>580
1-Hydroxy-1,1-ethane-diphosphonic acid, xNa Salt	29329-71-3	'low'	'low'	Re-testing 2Na salt	Nutrient complexation	Nutrient complexation	'low'
1-Hydroxy-1,1-ethane-diphosphonic acid, Na Salt	17721-68-5	'low'	'low'	Re-testing 2Na salt	Nutrient complexation	Nutrient complexation	'low'
1-Hydroxy-1,1-ethane-diphosphonic acid, 2Na Salt	7414-83-7	360	500	0.1	Nutrient complexation	3- (14d)	960
1-Hydroxy-1,1-ethane-diphosphonic acid, 3Na Salt	2666-14-0	'low'	'low''	Re-testing 2Na salt	Nutrient complexation	Nutrient complexation	'low'
1-Hydroxy-1,1-ethane-diphosphonic acid, 4Na Salt	3794-83-0	'low'	'low'	Re-testing 2Na salt	Nutrient complexation	Nutrient complexation	'low'
1-Hydroxy-1,1-ethane-diphosphonic acid, 5Na Salt	13710-39-9	'low'	'low'	Re-testing 2Na salt	Nutrient complexation	Nutrient complexation	'low'
1-Hydroxy-1,1-ethane-diphosphonic acid, xK Salt	67953-76-8	'low'	'low'	Re-testing 2Na salt	Nutrient complexation	Nutrient complexation	'low'
1-Hydroxy-1,1-ethane-diphosphonic acid, K Salt	17721-72-1	'low'	'low'	Re-testing 2Na salt	Nutrient complexation	Nutrient complexation	'low'
1-Hydroxy-1,1-ethane-diphosphonic acid, 2K Salt	21089-06-5	'low'	'low'	Re-testing 2Na salt	Nutrient complexation	Nutrient complexation	'low'
1-Hydroxy-1,1-ethane-diphosphonic acid, 3K Salt	60376-08-1	'low'	'low'	Re-testing 2Na salt	Nutrient complexation	Nutrient complexation	'low'
1-Hydroxy-1,1-ethane-diphosphonic acid, 4K Salt	14860-53-8	'low'	'low'	Re-testing 2Na salt	Nutrient complexation	Nutrient complexation	'low'

Substance	**CAS**	**Fish acute toxicity** **96h LC50 mg/l**	**Daphnia acute toxicity** **48h EC50 mg/l**	**Daphnia chronic toxicity** **22d NOEC mg/l**	**Algal toxicity** **96h EC50**	**Algal toxicity** **NOEC mg/l**	**toxicity to microorganisms** **30-min EC0 mg/l**
1-Hydroxy-1,1-ethane-diphosphonic acid, 5K Salt	87977-58-0	'low'	'low'	Re-testing 2Na salt	Nutrient complexation	Nutrient complexation	'low'

REFERENCES

Aptula A.O., Patlewicz G., Roberts D.W. & Schultz T.W. (2006). Non-enzymatic glutathione reactivity and in vitro toxicity: A non-animal approach to skin sensitization. *Toxicology In Vitro* **20**, 239-47.

ATC (2000a). An internationally recognised nomenclature system for petroleum additives. ATC Doc 31, rev 4. Technical Committee of Petroleum Additive Manufacturers in Europe.

ATC (2000b). Classification and user labelling information concerning the health effects of major petroleum additive components. ATC Doc 43, rev 4. Technical Committee of Petroleum Additive Manufacturers in Europe.

Benigni R, Conti L, Crebelli R, Rodomonte A & Vari' MR (2005). Simple and alpha, beta-unsaturated aldehydes: correct prediction of genotoxic activity through structure-activity relationship models. *Environ Mol Mutagen*. **46**, 268-280

Caley J, Henry T, Merckel D, Sonowane M, Benson A (2007). Generic lessons learnt from a review of categories presented at OECD Screening Information Assessment Meetings (SIAM) since April 2004. In: A Compendium of Case Studies that helped to shape the REACH Guidance on Chemical Categories and Read Across. Worth A & Patlewicz G (Eds). EUR report no 22481 EN. Available from ECB website: http://ecb.jrc.it/qsar/background/index.php?c=CAT.

CESIO (2000). CESIO recommendations for anionic and non-ionic surfactants. CEFIC, Brussels.

CESIO (2003). CESIO recommendation for the classification and labelling of surfactants as 'dangerous for the environment'. CEFIC, Brussels.

CONCAWE (2001). Classification and labelling of petroleum substances according to the EU dangerous substances directive (Recommendation, August 2001). Concawe, Brussels.

Clemedson C, Nordin-Andersson M, Bjerregaard HF, Clausen J, Forsby A, Gustafsson H, Hansson U, Isomaa B, Jorgensen C, Kolman A, Kotova N, Krause G, Kristen U, Kurppa K, Romert L & Scheers E. (2002). Development of an in vitro test battery for the estimation of acute human systemic toxicity: An outline of the EDIT project. Evaluation-guided Development of New In Vitro Test Batteries. *Alternatives to Laboratory Animals* **30**, 313-321.

Clemens, J.H., van den Heuvel, M.R., Stegeman, J.J., Dixon, D.G. and Bols, N.C. (1994) A comparison of toxic equivalency factors derived with rainbow trout, RTL-W1 and rat H4IIE cell lines for polychlorinated biphenyls (PCBs). *Society of Environmental Toxicology and Chemistry* Abstracts **14**, 280.

Comber M & Simpson B. (2007). Grouping of Petroleum substances. In: A Compendium of Case Studies that helped to shape the REACH Guidance on Chemical Categories and Read Across. Worth A & Patlewicz G (Eds). EUR report no 22481 EN. Available from ECB website: http://ecb.jrc.it/qsar/background/index.php?c=CAT.

CPIA (undated). Chlorinated Paraffins: A status report. Chlorinated Paraffins Industry Association. Accesssible from: http://www.regnet.com/cpia/publications.htm

ECB (2005). Scoping study on the development of a technical guidance document on information requirements on intrinsic properties of substances (RIP 3.3-1). Report prepared by CEFIC, DK-EPA, Environmental Agency of Wales and England, ECETOC, INERIS, KemI and TNO. European Chemicals Bureau, Joint Research Centre, European Commission, Ispra, Italy. Accessible from: http://ecb.jrc.it/DOCUMENTS/REACH/RIP_FINAL_REPORTS/RIP_3.3-1_INFO_REQUIREMENTS/

EC (2006a). European Union Risk Assessment Report, Cadmium oxide CAS-No.: 1306-19-0 & Cadmium metal CAS-No.: 7440-43-9, unpublished draft.

EC (2006b). European Union Risk Assessment Report, Nickel and Nickel Compounds, Nickel CAS-No.: 7440-02-0, Nickel Sulphate CAS-No.: 7786-81-4, Nickel Carbonate CAS No.: 3333-67-3; Nickel Chloride CAS-No.: 7718-54-9; Nickel Dinitrate CAS-No.: 13138-45-9, unpublished draft.

EC (2006c). European Union Risk Assessment Report, Pitch, coal tar, high-temp, CAS-No 65996-93-2, unpublished draft.

EC (2007). IUCLID 5 End User Manual. http://iuclid.eu

Environment Canada 2003. Guidance Manual for the Categorization of Organic and Inorganic Substances on Canada's Domestic Substances List. Existing Substances Branch, Environment Canada, Gatineau, Canada, 124 p

ETAD (2001). Guidance for the user industry on the environmental hazard labelling of dyestuffs. Ecological and toxicological assocaiation of dyes and organic pigments manufacturers. Basel, Switzerland.

Eurobitume (2002). Guidelines on the Classification of Bitumens & Bituminous Preparations. Eurobitume, Brussels.

Fabjan, E, Hulzebos E, Mennes W & Piersma A. (2006). A Category Approach for Reproductive Effects of Phthalates. *Critical Reviews in Toxicology* **36**, 695-726.

Gallegos Saliner A et al. (2007). Review of examples of read-across/category approaches discussed by the TC C&L. In: A Compendium of Case Studies that helped to shape the REACH Guidance on Chemical Categories and Read Across. Worth A & Patlewicz G (Eds). EUR report no 22481 EN. Available from ECB website: http://ecb.jrc.it/qsar/background/index.php?c=CAT.

Hanway, R.H. & Evans, P.F. (2000). Read-across of toxicological data in the notification of new chemicals. *Toxicology Letters* 116(S1), 61.

Hart J (2007). Nickel compounds – a category approach for metals in EU Legislation. A report to the Danish Environmental Protection Agency. http://cms.mim.dk/NR/rdonlyres/07DB028E-134E-4796-BF6D-97B9AD5F9E82/0/Nikkel.pdf

Hart J & Veith GD (2007): Applying Chemical Categories to Classification & Labelling: A Case Study of Volatile Aliphatic Ethers. A report to the Danish Environmental Protection Agency. Available at www.qsari.org/documents/aliphaticethers.pdf

HSPA (2002). The classification of petroleum sovent streams and related complex hydrocarbon solvents for aquatic environmental effects under the EU dangerous substances directive. Hydrocarbon Solvents Producers Association. Brussels 2002.

IHSC (2004/2005). Proposed hydrocarbon solvent category approach for the OECD SIDS programme. Presented at the January 2004 OECD Workshop on the development and use of chemical categories in the HPV Chemicals Programme, Workshop report, OECD Document ENV/JM/EXCH(2004)1.

Jäckh R (2007). Investigations of read-across & grouping, including mechanistic considerations and rationales for outliers. In: A Compendium of Case Studies that helped to shape the REACH Guidance on Chemical Categories and Read Across. Worth A & Patlewicz G (Eds). EUR report no 22481 EN. Available from ECB website: http://ecb.jrc.it/qsar/background/index.php?c=CAT.

OECD (2000). Guidance Document on Aquatic Toxicity Testing of Difficult Substances and Mixtures No. 23, ENV/JM/MONO(2000)6.

OECD (2001). OECD Environment, Health and Safety Publications. Series on Testing and Assessment, No. 29. Guidance Document on Transformation/Dissolution of Metals and Metal Compounds in Aqueous Media.

OECD (2004). OECD Principles for the Validation, for Regulatory Purposes, of (Quantitative) Structure-Activity Relationship Models
[http://www.oecd.org/document/23/0,2340,en_2649_34379_33957015_1_1_1_1,00.html]

OECD (2006a). Report on the Regulatory Uses and Applications in OECD Member Countries of (Q)SAR Models in the Assessment of New and Existing Chemicals. OECD Environment Health and Safety Publications. Series on Testing and Assessment. No. 58

OECD (2006b). OECD Harmonised Templates
[http://www.oecd.org/document/13/0,2340,en_2649_34379_36206733_1_1_1_1,00.html]

OECD (2007a). Guidance Document on the Validation of (Quantitative) Structure-Activity Relationships [(Q)SAR] Models. OECD Environment Health and Safety Publications. Series on Testing and Assessment. No. 69

OECD (2007b). Manual for Investigation of HPV Chemicals. http://www.oecd.org/document/7/0,2340,en_2649_34379_1947463_1_1_1_1,00.html

Parrott, J.L., Hodson, P.V., Dixon, D.G. and Servos, M.R. (1992) Liver EROD induction by several 2,3,7,8- chlorinated dibenzofurans in rainbow trout, *Society of Environmental Toxicology and Chemistry Abstracts* **13**, 113.

Salvito, D. (2007). Approaches to Chemical Categorization: An Illustrative Example of Approaches Used by the Fragrance Industry. In: A Compendium of Case Studies that helped to shape the REACH Guidance on Chemical Categories and Read Across. Worth A & Patlewicz G (Eds). EUR report no 22481 EN. Available from ECB website: http://ecb.jrc.it/qsar/background/index.php?c=CAT.

Safe, S. (1991) Polychlorinated dibenzo-*p*-dioxins and related compounds. Sources, environmental distribution and risk assessment, *Environmental Carcinogenicity and Ecotox. Reviews* **C9**(2), 261-302.

SCHER (Scientific Committee on Health and Environmental Risks) (2006): Opinion on Classification of Musk ketone. Adopted by the SCHER during the 9th plenary of 27 January 2006; scher_o_022.

Schoeters I & Verougstraete V (2007). Case studies on inorganic metal compounds. In: A Compendium of Case Studies that helped to shape the REACH Guidance on Chemical Categories and Read Across. Worth A & Patlewicz G (Eds). EUR report no 22481 EN. Available from ECB website: http://ecb.jrc.it/qsar/background/index.php?c=CAT.

Schultz, T.W., Cronin, M.T.D., Netzeva, T.I. & Aptula, A.O. (2002). Structure-toxicity relationships for aliphatic chemicals evaluated with Tetrahymena pyriformis. *Chemical Research in Toxicology* **15**, 1602-1609.

Tsakovska I & Worth A (2007). Review of examples of read-across/category approaches discussed by the TC NES. In: A Compendium of Case Studies that helped to shape the REACH Guidance on Chemical Categories and Read Across. Worth A & Patlewicz G (Eds). EUR report no 22481 EN. Available from ECB website: http://ecb.jrc.it/qsar/background/index.php?c=CAT.

UNEP Chemicals (2006) Screening Information Datasets for High Volume Chemicals, http://www.chem.unep.ch/irptc/sids/OECDSIDS/sidspub.html

US EPA (2004). Estimation of relative bioavailability of lead in soil and soil-like materials using in vivo and in vitro methods. OSWER 9285.7-77, Office of Solid Waste and Emergency Response, Washington, DC.

UV/EB Acrylate Resins (2003). Guide to the classification and labelling of UV/EB Acrylates. CEFIC, Brussels 2003.

van der Weiden, M.E.J., van der Kolk, J., Bleumink, R., Seinen, W. and van den Berg, M. (1992) Concurrence of P450 1A1 induction and toxic effects after administration of a low dose of 2,3,7,8-tetrachloro-*p*-dioxin (TCDD) in rainbow trout (*Oncorhynchus mykiss*), *Aquatic Toxicology* **24**, 123-142.

Walker, M.K. and Peterson, R.E. (1991) Potencies of polychlorinated dibenzo-*p*-dioxin, dibenzofuran and biphenyl congeners, relative to 2,3,7,8-tetrachloro-p-dioxin, for producing early life stage mortality in rainbow trout (*Oncorhynchus mykiss*), *Aquatic Toxicology* **21**, 219-238.

Worth A et al. (2007). Possible applications of QSAR methods. In: A Compendium of Case Studies that helped to shape the REACH Guidance on Chemical Categories and Read Across. Worth A & Patlewicz G (Eds). EUR report no 22481 EN. Available from ECB website: http://ecb.jrc.it/qsar/background/index.php?c=CAT.

Zabel, E.W., Cook, P.M. and Peterson, R.E. (1995) Toxic equivalency factors of polychlorinated dibenzo-*p*-dioxin, dibenzofuran and biphenyl congeners based on early life stage mortality in rainbow trout (*Oncorhynchus mykiss*), *Aquatic Toxicology* **31**, 315-328.

LIST OF ABBREVIATIONS

AD	Applicability Domain
BCF	Bioconcentration Factor
CAS	Chemical Abstracts Service
CEFIC	European Chemical Industry Council
CESIO	Comité Européen des agents de Surface et de leurs Intermédiares Organiques
CPIA	Chlorinated Paraffins Industry Association
CONCAWE	The oil companies' European Organisation for Environment, Health and Safety in Refining and Distribution
DNEL	Derived No Effect Level
EPA	Environmental Protection Agency
ESR	Existing Substances Regulation (European Union)
ETBE	ethyl tert-butyl ether
EU	European Union
EWG	Endpoint working group
GHS	Globally Harmonised System (for the classification of chemicals)
HPV	High Production Volume
HSDB	Hazardous Substances Database
HSPA	Hydrocarbon Solvents Producers Association
IHSC	International Hydrocarbon Solvents Consortium
ITS	Intelligent Testing Strategy
IUCLID	International Uniform ChemicaL Information Database
IWG	Information Working Groups
Kow	Octanol-water partition coefficient
logKow	log of the octanol-water partition coefficient
LC50	Concentration of a compound that causes 50% lethality of the animals in a test batch
LD50	Dose of a compound that causes 50% lethality of the animals in a test batch
MCS	Multi-constituent substance
MTBE	methyl tert-butyl ether
MW	Molecular Weight

NCS	Natural Complex Substances
NGO	Non Governmental Organisation
NLM	National Library of Medicine (USA)
NOAEL	No Observable Adverse Effect Level
NOEC	No Observed Effect Concentration
NONS	Notification of New Chemicals (European Union)
OECD	Organization for Economic Cooperation and Development
PAH	Polyaromatic Hydrocarbon
PBT	Persistent, Bioaccumulative and Toxic
PMG	Project Management Group
PNEC	Predicted No Effect Concentration
QAAR	Quantitative Activity-Activity Relationship
QSAR	Quantitative Structure-Activity Relationships
RCR	Risk Characterisation Ratio
REACH	Registration, Evaluation, Authorisation of Chemicals (European Union)
RIP	REACH Implementation Project (European Union)
SAR	Structure Activity Relationship
SCHER	Scientific Committee on Health and Environmental Risks (European Union)
SIAM	SIDS Initial Assessment Meeting (OECD)
SIAR	SIDS Initial Assessment Report (OECD)
SIDS	Screening Information Data Set (OECD)
SMILES	Simplified Molecular Input Line Entry System
TAME	tert-amyl methyl ether
TAPIR	Three point three – A Project for the Information requirements of REACH
TC C&L	Technical Committee for Classification and Labelling (European Union)
TCEP	Tris(2-chloroethyl) phosphate
TC NES	Technical Committee on New and Existing Substances (European Union)
TCPP	Tris(2-chloro-1-methylethyl) phosphate
TDCP	Tris[2-chloro-1-(chloromethyl)ethyl] phosphate
TEF	Toxic Equivalency Factor
TEQ	Toxic Equivalents (Approach)
TGD	Technical Guidance Document
TOXNET	Toxicology Data Network
UNEP	United Nations Environment Programme

UVCB	Substances of Unknown or Variable composition, Complex reaction product or Biological material
vPvB	Very Persistent and Very Bioaccumulative
WoE	Weight of Evidence approach