Unclassified **ENV/JM/MONO(2014)4**

Organisation de Coopération et de Développement Économiques
Organisation for Economic Co-operation and Development **14-Apr-2014**

English - Or. English

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

GUIDANCE ON GROUPING OF CHEMICALS, SECOND EDITION

Series on Testing & Assessment
No. 194

JT03356214

Complete document available on OLIS in its original format
This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

ENV/JM/MONO(2014)4
Unclassified
English - Or. English

OECD Environment, Health and Safety Publications
Series on Testing and Assessment

No. 194

GUIDANCE ON GROUPING OF CHEMICALS, SECOND EDITION

INTER-ORGANIZATION PROGRAMME FOR THE SOUND MANAGEMENT OF CHEMICALS
A cooperative agreement among FAO, ILO, UNDP, UNEP, UNIDO, UNITAR, WHO, World Bank and OECD

Environment Directorate
ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT
Paris 2014

About the OECD

The Organisation for Economic Co-operation and Development (OECD) is an intergovernmental organisation in which representatives of 34 industrialised countries in North and South 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 11 different series: **Testing and Assessment; Good Laboratory Practice and Compliance Monitoring; Pesticides; 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 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 (www.oecd.org/chemicalsafety/).

This publication was developed in the IOMC context. The contents do not necessarily reflect the views or stated policies of individual IOMC Participating Organisations.

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, UNDP, UNEP, UNIDO, UNITAR, WHO, World Bank and OECD. 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.

Also published in the Series on Testing and Assessment link

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

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

© OECD 2014
Applications for permission to reproduce or translate all or part of this material should be made to: Head of Publications Service, RIGHTS@oecd.org, OECD, 2 rue André-Pascal, 75775 Paris Cedex 16, France

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. Instead, the overall data for that category should prove adequate to support a hazard assessment. The overall data set must enable an estimate of hazard for the untested endpoints.

Although this approach has been used on an *ad hoc* basis in many regulatory programmes for many years, guidance was first developed by the US-EPA in support of the US HPV Challenge Program in 1998. The same guidance was also embedded into the *OECD Manual for the Assessment of Chemicals*. Since then, guidance has evolved continuously based on experience with the approach within the OECD Cooperative Chemicals Assessment 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 recommending its wider use. Since the technique of assessing groups of substances is an evolving science, this guidance document is revised periodically. Furthermore, due to the developing nature of the approach as well as its complexity, early consultations between industry and authorities are recommended to ensure that any regulatory requirements are fulfilled if applying it for that purpose. The OECD Guidance reflects on the elements common to a number of different applications and real examples of grouping approaches to help users understand basic concepts. Users will need to take account of the fact that the OECD Guidance cannot cover all the regulatory requirements that may apply to their situation. The present document is the second edition of the guidance, initially published in 2007. This edition has been augmented with experience and examples encountered in the OECD Cooperative Chemicals Assessment Programme, formerly the HPV Chemicals Programme since 2007. The second edition also intends to introduce new or revised guidance on: elaborating the analogue and category approach, quantitative and qualitative read-across, justifying read-across, using bioprofiling results for grouping chemicals, and specific types of category approaches (e.g. chemicals of variable composition, and metals).

The guidance first explains in Chapter 2 what a category is and outlines relevant concepts that will enable the reader to better understand the remainder of the document. This chapter outlines general aspects of grouping chemicals such as the identification of analogues /members of categories, the mechanistic basis for using analogues or chemical categories, and the robustness of both approaches. Chapter 2 also describes the close relationship that exists between (Q)SARs and categories, both in terms of the concepts and in the use of (Q)SARs for data evaluation and data gap filling. Chapter 3 explains the main approaches that are used for data gap filling: read-across, trend analysis and (Q)SARs. While Chapters 2 and 3 provide explanations on the scientific and methodological background of the analogue and category approaches, respectively, Chapters 4-7 focus on practical aspects for forming and documenting analogue and chemical category approaches. Separate chapters (4 and 5) were elaborated to provide guidance on the stepwise procedures for analogue and chemical category approaches, such that the guidance document can be used in a "modular" fashion, and therefore making it possible to use parts of the guidance only. Accordingly a number of text repetitions were necessary. Chapter 6 elaborates on some of the specific issues that need to be addressed with certain types of chemical substances. Finally, in Chapter 7, formats are proposed to structure the documentation of analogue and category approaches, so-named analogue/category reporting formats (ARF, CRF).

This document is published under the responsibility of the Joint Meeting of the Chemicals Committee and the Working Party on Chemicals, Pesticides and Biotechnology.

TABLE OF CONTENTS

1. INTRODUCTION 9

2. EXPLANATION OF THE GROUPING APPROACHES 11

2.1. Introduction and concepts 11
2.2. Considerations when grouping chemicals 14
2.3. Selecting analogues/Creating chemical categories and setting boundaries 16
2.3.1. Selecting analogues 16
2.3.2. Category and subcategory membership and applicability domain 16
2.4. The mechanistic basis of using analogues or chemical categories 19
2.4.1. Principle and considerations 19
2.4.2. The concept of Adverse Outcome Pathways and the use of bioprofiling information for grouping chemicals 21
2.5. Robustness of a chemical category and of an analogue approach 22
2.6. The interdependence between categories and (Q)SARs 24

3. TECHNIQUES OR METHODS FOR DATA GAP FILLING 26

3.1. Introduction 26
3.2. Read-across 27
3.2.1 Hypothesis and evidence based approaches 28
3.2.2 Choice of qualitative or quantitative read-across 28
3.2.3 Elements for a read-across justification 32
3.3. Trend analysis and computational methods based on internal models 52
3.3.1 Examples of trend analysis and breakpoints 54
3.4. Computational methods based on external models 55
3.5. Reporting on incremental changes, trend analysis and computational methods 56

4. ANALOGUE APPROACH: A STEPWISE PROCEDURE FOR IDENTIFYING ANALOGUES AND READ-ACROSS 58

4.1. Introduction 58
4.2. Stepwise approach to read-across using the analogue approach 59
4.2.1. Step 0: Check whether the chemical is a member of an existing category 59
4.2.2. Step 1: Identification of potential analogues 59
4.2.3. Step 2: Data gathering for the analogues 61
4.2.4. Step 3: Evaluation of available data for adequacy 62
4.2.5. Step 4: Construct a matrix of data availability 62
4.2.6. Step 5: Assess the adequacy of the analogue approach and fill the data gap 62
4.2.7. Step 6: Document the analogue approach 63

5. CATEGORY APPROACH: A STEPWISE PROCEDURE FOR GROUPING CHEMICALS AND READ-ACROSS 66

5.1. Introduction 66
5.2. Stepwise approach to the formation of chemical categories 67
5.2.1. Step 0: Check whether the chemical is a member of an existing category 67
5.2.2. Step 1: Develop category hypothesis and definition and identify category members 68
5.2.3. Step 2: Gather data for each category member 71
5.2.4. Step 3: Evaluate available data for adequacy 71

5.2.5. Step 4: Construct a matrix of data availability 74
5.2.6. Step 5: Perform a preliminary evaluation of the category and fill data gaps 74
5.2.7. Step 6: Perform and/or propose testing 75
5.2.8. Step 7: Perform a further evaluation of the category 76
5.2.9. Step 8: Document the finalised category and refine the category rationale 76

6. GUIDANCE ON SPECIFIC TYPES OF CATEGORIES 78

6.1. Chain length 78
6.2. Metabolic or degradation pathways and toxicokinetics 79
6.3. Mixtures of discrete substances 82
6.4. Use of toxic equivalency factors or toxic units approach for filling data gaps 83
6.5. Isomers 85
6.6. Complex substances (UVCBs) 87
6.6.1. General guidance on developing categories for organic UVCBs 88
6.6.2. Hydrocarbon solvents 93
6.6.3. Coal derived complex substances 94
6.6.4. Natural complex substances (NCS) 94
6.6.5. Developing categories for complex inorganic UVCB substances 95
6.7. Metals and inorganic metal compounds 96
6.7.1. Introduction 96
6.7.2. Basis and assumptions underlying the grouping of metal compounds 97
6.7.3. Grouping of metals: in practice 99
6.7.4. Need to verify the assumptions and further refinements 102
6.8. General guidance for other compounds (e.g. ionisable compounds) 103
6.9. Initial considerations applicable to manufactured nanomaterials 103

7. REPORTING FORMATS FOR ANALOGUE AND CATEGORY EVALUATIONS 105

7.1. Reporting Format for analogue approach 105
7.2. Reporting format for chemical categories 109

8. REFERENCES 115

9. LIST OF ABBREVIATIONS 126

10. APPENDIX 128

1 INTRODUCTION

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

The practice of predicting properties of chemicals is already established in regulatory science, and improved techniques are evolving as scientific knowledge develops and is applied in this field. The OECD Guidance on grouping of chemicals tries to encompass different possibilities and interpretations. The way in which grouping is undertaken to predict properties of some members of the group depends on the purpose of the prediction, e.g., for commercial decision-making, screening and priority-setting of chemicals for further evaluation, hazard identification for risk assessment and classification and labelling, filling information requirements in different regulatory schemes. Therefore, the administrative practice, standard of proof, and degree of scientific certainty in the assessment will all vary depending on the purpose of the prediction.

For reasons of resources and animal welfare, it is important to reduce as much as possible the number of *in vivo* tests to be conducted where scientifically justifiable. One approach is to consider closely related chemicals as a group rather than as individual chemicals. If grouping is applied, not every chemical needs to be tested for every required endpoint. Rather, the data for chemicals and endpoints that have been tested can be used to estimate the corresponding properties for the untested chemicals and endpoints. Grouping of chemicals can lead to the application of a category or an analogue approach.

An 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 chemicals in the category, compared to evaluation of data purely on a chemical-by-chemical basis.

In the analogue approach 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).

All category assessments should be reviewed and updated, when new information is generated because category assessments are often complex and experience in forming and assessing categories is continuously growing. Periodic review and update of category assessments provides a means of incorporating new information, re-affirming or strengthening the scientific basis of the original hypothesis 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 that calls the category justification into question. In such cases, the category should be re-evaluated and may need to be re-constructed.

This document has been developed based on existing cases involving chemical categories assessed within the OECD Cooperative Chemicals Assessment Programme[1] (formerly the OECD HPV Chemicals

[1]OECD Cooperative Chemicals Assessment Programme (CoCAP) http://www.oecd.org/env/ehs/risk-assessment/oecdcooperativechemicalsassessmentprogramme.htm

Programme), the US HPV Challenge Program[2], the EU Existing Substances Programme (replaced by REACH[3] in 2009), the EU activity on classification and labelling[4], Canada's Chemicals Management Plan[5] (CMP) and Domestic Substances List (DSL) program[6], guidance issued under the US HPV Challenge Program and other US EPA programs as well as for the EU REACH legislation, and the experience gained from the OECD Workshop on the development and use of chemical categories held in January 2004. This document was updated in 2013 to take into consideration progress made in the mechanistic understanding of interactions between the chemical and the biological target and key events leading to adverse effects in whole organisms (i.e., the Adverse Outcome Pathway (AOP) concept). This update also reflects experience gained from the OECD Workshop on Using Mechanistic Information in Forming Chemical Categories (OECD, 2011a). The document addresses the actual formation of categories for test plan and hazard assessment purposes. It also provides guidance about how to present data, e.g., data matrices showing all data available for category members (both standard test data and data from alternative methods, with an indication of the data gap filling technique that is proposed.

"Data gap filling" is the process of providing data to inform upon a particular endpoint by whatever means is scientifically justified including alternative techniques to direct testing. Read-across and trend analysis are methods that may be used for data gap filling as described in this document.

The regulatory application of (Q)SAR 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 the Assessment of Chemicals provides guidance on the use of SAR in the HPV Chemicals Programme (OECD 2011b);
- 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 (Q)SAR applications;
- OECD report on the principles for the validation, for regulatory purposes, of (Q)SAR models (OECD 2004a) and an accompanying OECD guidance document (OECD 2007a);
- OECD Guidance Document for Using the OECD (Q)SAR Application Toolbox to Develop Chemical Categories According to the OECD Guidance on Grouping Chemicals (OECD, 2009a);
- Report of the Workshop on Structural Alerts for the OECD (Q)SAR Application Toolbox (OECD, 2009b);
- Training for the QSAR Toolbox; (http://www.oecd.org/fr/env/ess/risques/guidancedocumentsandreportsrelatedtoqsars.htm)
- NAFTA (2012). Technical working group on pesticides (TWG) (Quantitative) structure activity relationship [(Q)SAR] guidance Document. http://www.epa.gov/oppfead1/international/naftatwg/guidance/qsar-guidance.pdf

[2] US EPA High Production Volume (HPV) Challenge Program http://www.epa.gov/hpv/

[3] REGULATION (EC) No 1907/2006 of The European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH),establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC

[4] REGULATION (EC) No 1272/2008 of The European Parliament and of the Council of 16 December 2008on classification, labelling and packaging of substances and mixtures, amendin and repealing Directives 67/548/EEC and 1999/45/EC, and amending Regulation (EC) No 1907/2006

[5] Canada - Chemicals Management Plan http://www.chemicalsubstanceschimiques.gc.ca/plan/index-eng.php

[6] Environment Canada - Domestic Substances List http://ec.gc.ca/lcpe-cepa/default.asp?lang=En&n=5F213FA8-1

2. EXPLANATION OF THE GROUPING APPROACHES

2.1 Introduction and concepts

In this OECD guidance document, the term 'grouping' or 'chemical grouping' describes the general approach for considering more than one chemical at the same time. It can include formation of a chemical category or identification of (a) chemical analogue(s) with the aim of filling data gaps[7] as appropriate. The category or the analogue approach makes it 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 this way, both approaches are important since they provide an alternative to testing individual chemicals and as a result should lead to a decrease in the use of animal testing. In addition it will increase the knowledge of the hazard properties of chemicals that may otherwise remain untested and provide for an increased level of protection for human health and the environment.

Analogue approach

When the focus of the assessment is on filling data gaps for one specific chemical, empirical data from one or more similar chemical(s) ("the analogue(s)") [8] or "source" chemical can be used to predict the same endpoint[9] for the "target" chemical[10,] which is considered to be "similar." This analogue approach is useful when the target and source chemicals share a known common mode (and/or mechanism) of action[11 12,] and the adverse effects[13] resulting from this mode (and/or mechanism) of action is evaluated. The analogue approach could also be used in the absence of effects or when no specific mode (and/or mechanism) of action is expected and toxicokinetic behaviour is not expected to differ significantly. In such case, more evidence,[14] or more lines of evidence, should support the assessment.

Category approach

Chemicals whose physical-chemical, toxicological and ecotoxicological properties are likely to be similar or follow a regular pattern as a result of structural similarity may be considered as a group, or 'category' of

[7] A data gap is a physical-chemical, environmental fate, ecotoxicological, or mammalian toxicological/human health endpoint for which data are not available when required for an assessment.
[8] An analogue is a chemical whose intrinsic physical-chemical, environmental or toxicological properties are likely to be similar to another chemical based upon a number of potential properties, including structural, physical-chemical and toxicological.
[9] An endpoint refers to a broad description of a specific environmental or toxicological property, for example acute oral toxicity, or water solubility.
[10] A target chemical is one with data gap(s), for which a property or hazard is being estimated from the source chemical(s).
[11] A mode of action describes a functional or anatomical change, at the cellular level, resulting from the exposure of a living organism to a chemical. In comparison, a mechanism of action describes such changes at the molecular level.
[12] A mechanism of action denotes the sequence of events leading from the absorption of an effective dose of a chemical to the production of a specific biological response in the target organ. Understanding a chemical's mechanism requires appreciation of the causality and temporal relationships between the steps leading to a particular toxic endpoint, as well as the steps that lead to an effective dose of the chemical at the relevant biological target(s).
[13] An adverse effect refers to the change in the morphology, physiology, growth, development, reproduction, or life span of an organism, system, or (sub)population that results in an impairment of functional capacity, an impairment of the capacity to compensate for additional stress, or an increase in susceptibility to other influences (ICPS, 2004).
[14] In the context of grouping chemicals, evidence refers to similarities in chemical data that is used to justify reading across from source to target chemical(s) and developing chemical categories.

chemicals. The assessment of chemicals by using this category approach differs from the approach of assessing them on an individual basis, since the properties 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 chemical alone. For (a) category member(s) that lacks data for one or more endpoints, the data gap can be filled in a number of ways, including by read-across from one or more other category members. Within a chemical category, the members are often related by a trend in an effect for a given endpoint, and a trend analysis[15] can be carried out through deriving a model based on the data for the members of the category.

The rationale underpinning the analogue and the category approach may be based on the following:

- Common functional group(s) (e.g., aldehyde, epoxide, ester, specific metal ion);
- A common mode or mechanism of action or adverse outcome pathway
- Common constituents or chemical classes, similar carbon range numbers. This is frequently the case with complex substances[16] often known as "substances of unknown or variable composition, complex reaction products or biological material" (UVCB substances);
- The likelihood of common precursors and/or breakdown products via physical or biological processes that result in structurally similar chemicals (e.g., the "metabolic pathway approach" of examining related chemicals such as acid/ester/salt); or
- An incremental and constant change across the category (e.g., a chain-length category), often observed in physical chemical properties, e.g., boiling point range.

For every category the structural elements that category members have in common need to be described together with structural differences that may occur in the category. These differences may or may not affect the endpoint of interest. Differences, which are not expected to affect the endpoint of interest could be called "allowed differences." The category could then further be limited by other criteria not related to the structure of the category members. While a category may in principle be based on one of these rationales, in practice endpoint justifications and supporting information will be multifaceted. All pre-existing experimental or other (e.g., from the literature) evidence that can support the category needs to be addressed. This could be similar effects in lower-tier studies where these exist, availability of "bridging" studies[17] that are not necessarily endpoint related (e.g., common results in in vitro or other types of screening studies), evidence from computational and non-computational theoretical models, common bioavailability[18], metabolism[19] and reactivity profiles, common mode and/or mechanism of action (MOA), or adverse outcome pathway (AOP)[20].

[15] Trend analysis refers to a data-gap filling method for "quantitative endpoints" (e.g., 96h-LC50 for fish) if a number of analogues (at least 3) with experimental results are identified.

[16] A substance is defined as a chemical element and its compounds in the natural state or obtained by any manufacturing process, including any additive necessary to preserve its stability and any impurity deriving from the process used, but excluding any solvent which may be separated without affecting the stability of the substance or changing its composition. A substance may contain one or more main constituents (i.e., constituent(s) that make(s) up a significant part of that substance). The main constituent(s) should clearly be other than impurities (i.e., all the unintentional constituents coming from the manufacturing process or from the starting material(s); these could be the result of secondary or incomplete reactions occurring during the production and are present in the final substance even if not sought by the manufacturer) and additives (i.e., all the constituents which are intentionally added to stabilise the substance and only for this purpose).

[17] Bridging studies are defined but not limited to studies conducted to show relevance or create a bridge between existing studies to avoid replicating existing studies.

[18] Bioavailability is defined as the extent to which a substance is taken up by an organism and distributed to an area within the organism. It is dependent upon physical-chemical properties of the substance, anatomy and physiology of the organism, pharmacokinetics, and route of exposure. Availability is not a prerequisite for bioavailability. (United Nations, 2013). An alternative definition for bioavailability is the rate and extent to which a substance can be taken up by an organism and is

Presentation of experimental results and presentation of the data gap filling approach used can be facilitated by the use of a data matrix. Examples of the data matrices used to report the use of this approach are shown in Chapter 7. Further guidance on the justification of data gap filling is given in Chapter 3. General guidance on how to apply and justify the analogue and category approaches is provided in Chapters 4 and 5, respectively. Specific guidance for different types of categories is given in Chapter 6.

While other data gap filling approaches will be discussed in more detail in Chapter 3, it is worth mentioning interpolation or extrapolation. Within a category where trends in toxicity or factors influencing toxicity have been identified and the category members arranged in line with the trend as illustrated in Figure 1, interpolation can be described as the process whereby data from category members on either side of a data-poor category member is used to predict its hazards. In contrast, extrapolation is the process where data from category members at one side of the category is used to predict the hazards of those members at the other side. Of course, it could also be said that an analogue approach itself is by default an extrapolation, unless there are analogues identified that bracket the target chemical.

There is a preference for the use of interpolation rather than extrapolation, because extrapolation is perceived to be more uncertain and therefore less reliable. The quality of the trend can influence the uncertainty, among other factors. Therefore, although it may seem logical for interpolation to be more 'acceptable' than extrapolation, the degree of uncertainty is not really due to the interpolation or extrapolation of data, but rather the robustness of the category. Robustness is in turn dependent on the size of the category and the amount of data available for each category member. If a trend is poorly defined or missing, then interpolation and extrapolation approaches will be equally uncertain. The establishment of a trend requires two variables: dependent and independent. Thus, the trend will depend also on the choice and quality of the independent variable.

In large, data-rich categories, trends in toxicity are more likely to be readily characterised, such that any data gap filling, whether it be interpolation or extrapolation, is more likely to be robust and useful. However, in cases where an analogue approach has been used or where a category consists of only a small number of members, trends can be far more difficult to identify and interpolation is often not possible simply due to the small number of members. In these situations, extrapolation of data from one chemical to another may be the only possibility for filling the data gaps. In order to minimise uncertainty, a weight of evidence (WoE)[21] proposal can be developed that incorporates the use of extrapolation together with information from (Q)SAR tools, in vitro assays or other mechanistic/bridging studies. There will always be some degree of uncertainty, but this should be considered in light of the inherent uncertainties associated with all test data, and not just when applying the category or analogue approach (ECETOC, 2012). Nevertheless, the transfer of data from one or more substances to another, may bring extra uncertainty compared to the generated experimental data.

---

available for metabolism or interaction with biologically significant receptors. Bioavailability (biological availability) involves both release from a medium (if present) and absorption by an organism (IPCS 2004).

[19] Metabolism is the sum total of all physical and chemical processes that take place within an organism from uptake to elimination.

[20] Adverse outcome pathways delineate the documented, plausible, and testable processes by which a chemical induces molecular perturbations (Molecular Initiating Events) and the associated biological responses that describe how the molecular perturbations cause effects at the subcellular, cellular, tissue, organ, whole animal, and population levels of observation.

[21] Weight of evidence refers to a positive expert opinion that considers available evidence from different independent sources and scientific viewpoints on a particular issue, coming to a considered view of the available, oftentimes conflicting, data. It is preferred when every source does not provide sufficient information individually.

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

	Chemical 1	Chemical 2	Chemical 3	Chemical 4	
Structure	xxxxxxxxx	xxxxxxxxx	xxxxxxxxx	xxxxxxxxx	
Property 1	●	○	●	○	SAR/Read-across
Property 2	●	○	○	●	Interpolation
Property 3	○	●	●	○	Extrapolation
Activity 1	●	○	●	○	SAR/Read-across
Activity 2	●	○	○	●	Interpolation
Activity 3	○	●	●	○	Extrapolation

● Existing data point ○ Missing data point

2.2 Considerations when grouping chemicals

While there are many benefits to the use of analogue and category approaches, there are also some caveats to consider upfront. The benefits are outlined in detail below and aside from the scientific insights that can be potentially derived, these could also be categorised in terms of savings, i.e., money, time, and animals. The caveats or hurdles are both scientific and practical in nature. The practical hurdles are largely cost and procedure based, i.e., gaining access to good quality data that are required for the data gap filling approach, documenting the appropriate level of study information, and having the necessary information to characterise the target chemical and source analogues as well as their respective impurity profiles. The cost in this sense merely outlines the potentially significant upfront costs associated with gaining access to, or generating the data to be used to inform data gap filling or indeed the costs of new data that may need to be generated to help substantiate the hypothesis being used as the basis of the category/analogue approach in the first instance. Such costs should typically be lower than the costs associated with performing the experimental studies. Another practical consideration could be the number of data gaps that need to be filled for a specific chemical; an analogue/category approach could be conceivably waived in favour of the use of (Q)SAR approaches alone. Obviously, this will depend on the endpoints under consideration and the maturity of the (Q)SAR models available for use. More information on (Q)SARs is discussed in Chapter 3. Scientific hurdles exist, such as the level of mechanistic understanding required for specific endpoints to help inform the biological plausibility of grouping. Some endpoints such as skin sensitisation have been characterised using the mode of action/adverse outcome pathway (MOA/AOP) concept that facilitates building toxicologically meaningful categories, whereas other endpoints are less well understood, which raises the uncertainty of the data gap filling techniques. As such there is a potential risk in over- or under-characterising the hazards of a specific chemical under consideration.

Both the analogue and chemical category approach share the same benefits, in that data from one or more chemicals can be related to other chemicals, reducing the need to test every endpoint for every chemical. In addition, the 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:

- The identification of compounds as members of a category provides an insight into the potential effects of the compounds that might otherwise be overlooked.
- 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 these can present technical difficulties when carrying out standard test protocols (examples are given in Hart (2007) and Comber & Simpson (2007).
- In order to gain future efficiencies, category proposals may be expanded via the inclusions of chemicals that may be addressed under various global programmes.

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 category 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. 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 *in vitro* or 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. This approach is further elaborated in the concept of MOA/AOPs (see Section 2.4.2).

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 common mode (and/or mechanism) of action, the confidence in the category is significantly greater than that associated with the use of an analogue approach where the mode (and/or mechanism) of action of the target and source chemical is unknown. This confidence increases with increasing numbers of chemicals and with empirical data included in the category. For a large category[22], 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 category can be identified with the necessary confidence. For more limited comparisons, particularly with chemicals containing multiple functional groups, it may be harder to obtain the same level of confidence. For filling data gaps, a category approach can provide significant advantages compared to the analogue approach in that, with the category approach, it is possible to analyse trends in properties. Within an analogue approach, confidence might also be derived if both target and source chemicals display a consistent pattern across many endpoints. Data gap filling techniques between chemical analogues have been extensively used (e.g., within the OECD Cooperative Chemicals Assessment Programme, formerly the OECD HPV Chemicals Programme, the EU Existing Chemicals Programme or for Classification and Labelling in the EU, the U.S. EPA Voluntary High Production Volume (HPV) Challenge Program, and under the Chemicals Management Plan in Canada), often on an *ad hoc* basis and it is foreseen that these will continue to be used extensively. Nevertheless, an important consideration in revising this Guidance is to encourage the replacement of these *ad hoc* approaches with a more systematic approach that can provide a greater degree of transparency in the result.

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

2.3 Selecting analogues/Creating chemical categories and setting boundaries

2.3.1 Selecting analogues

There are a number of different ways of identifying potential analogues as source chemicals with data with which the target chemical can be compared. The approaches and tools are described in more detail in Chapter 4. In some cases, the choice of a source chemical may be straightforward; such as where similar chemicals are produced for similar uses by the same company (or sector group of companies). In this case, no formal identification techniques may be required. However, a more formal search strategy may identify additional analogues for comparison, and hence potentially increase the robustness of the subsequent data gap filling. Evaluation of analogues is a critical step. The rationales described in Section 2.1 are useful starting points to characterise the underlying hypothesis that will be used in support of a category or analogue approach. General considerations for evaluating analogues include an assessment of the physical-chemical, reactivity and metabolic similarity (ECETOC, 2012). These are discussed in more detail in Section 2.4. Tools such as the OECD QSAR Toolbox (OECD, 2013), which contains profilers based on rules defining the structural and referential boundaries, can be helpful to systematically assess analogues on these bases. These profilers can also be used to apply to new analogues as well as to set the boundaries for creating new categories.

2.3.2 Category and subcategory membership and applicability domain

1. Category membership and applicability domain

In an ideal situation, a category would identify all potential members of the category when first developed. A high-level grouping via clustering[23] of chemical inventories would facilitate such an exploration. However, this ideal situation is difficult to achieve in practice. Inclusion or exclusion of certain substances could introduce bias to the data gap filling. Therefore, the choice of category members should be always explained. 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 that are also included in the category. Therefore, a clear category definition and description should allow a category to be expanded with additional substances. The category definition includes the category name and category members (i.e., IUPAC chemical names, CAS numbers, and structures). The category description should include a summary of common features; boundaries; physical-chemical properties, if applicable; allowed variations in chemical structure; and if known, any restrictions (e.g., variations that would change the effects of a substance significantly compared to the other substances in the category). The identity of the test material, when experimental data are available, should be known.

Practical considerations will often influence the choice of chemicals included in the category. The selection of chemicals that are included in a particular chemical category is frequently guided by which chemicals are manufactured by the consortium of companies sponsoring the 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 these are produced by the companies carrying out the category evaluation. The practical considerations should not overwhelm the toxicological reasoning for grouping. Otherwise, these can introduce bias to the data gap filling.

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

[23] A cluster is a group of chemicals organised according to similar characteristics, such as structure.

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

If a chemical is assessed and subsequently identified as a potential member of an existing category, it may be necessary to evaluate both the data for this chemical in light of the category evaluation and the category evaluation in light of the data for the additional chemical. If the initial category evaluation is sufficiently robust, the additional data are unlikely to alter the conclusions of the initial evaluation, but additional data may strengthen the category further. 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 potential revisions resulting from adding data at a later date. Experience has shown that, in many cases, additional chemicals identified fall on either the lower or upper boundary[24] 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 judgment and WoE (See Chapter 3, Section 3.5) are used together in making recommendations/decisions about the level of testing that may be required, if any.

When assessing whether a chemical could be a member of an existing category (of which it is not already a member), clarity on the applicability domain (AD) of the category (i.e., which chemicals are covered by the category assessment) is important for the regulatory acceptance of the hazard conclusions. The applicability domain of a category would ideally identify the structural requirements and ranges of physicochemical, environmental fate, toxicological or ecotoxicological properties within which reliable estimations can be made for the category members. For this reason, the precise composition of the category (e.g., carbon number range, branching and position of branching, aromatic content, cyclicity, position and frequency of double bonds, functional group(s)) of category members) should be defined where possible to set the boundaries that are used as inclusion/exclusion criteria. The applicability domain boundaries would be also supported by the commonality in MOA/AOP – some of which may be encoded in structural features characterising molecular initiating events[25]. 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 acute 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.

Defining the AD is also important because later additions to the category will require reconsideration of the data gap filling approach and results. If certain endpoints do not follow a trend, care needs to be exercised to determine whether the category is still justified for those endpoints, i.e., whether and what type of techniques can be applied to fill a data gap.

2. Subcategories

In some cases, an effect can be present or follow a trend 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 while higher members do not (OECD, 2004b). For other properties/effect types, 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 & Veith 2007).

[24] Category members falling at the opposite extremes of a trend and within which interpolations are considered reliable are called sentinel or boundary chemicals (OECD, 2007).
[25] A molecular initiating event is a chemical interaction at a molecular target in an organism that leads to a particular adverse outcome.

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

- An effect that 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., these have degrees of hazard potential or different regulatory classifications) or quantitative (i.e., the numerical values of the endpoint include values on either side of a breakpoint[26]); or
- An effect where there is a peak in activity or a breakpoint in a trend can also lead to the formation of subcategories; or
- 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 this approach. The organisation of the category in subcategories should be presented and justified in the category justification.

Examples that have been encountered within the OECD Cooperative Chemicals Assessment Programme 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 (OECD 2004b). 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 *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 was 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 was reduced (OECD, 2006b).

3. *Categories for human health or for the environment*

Sometimes the category approach may be applicable and justified for human health endpoints (e.g., same functional group, same metabolism, or same mode (and/or mechanism) of action), but not for environmental endpoints (e.g., different environmental fate, different aquatic toxicity across members of the category), and vice-versa. An example includes the C2-C4 aliphatic thiols category where hazardous properties identified for human health are identical across the four members of the category (irritation, skin sensitization, repeated-dose toxicity (OECD, 2010a). Category members also share identical acute aquatic toxicity properties, but the environmental fate properties differ and result in different hazard conclusions. Another example could be a homologous series of alkanes; the aquatic toxicity would be expected to follow a trend based on chain length but differences could be expected for human health effects due to metabolism. Hexane and pentane are examples. Hexane's toxicity is mediated by its metabolite hexane-2,5-dione; whereas pentane is hydroxylated to its corresponding alcohol (ECETOC, 2012). When no additional information was available, these will be outliers in an otherwise homologous looking series. The appearance of outliers brings extra uncertainty to the prediction and their exclusion should be accompanied with appropriate understanding. A category approach could also be applicable to many human health effects where metabolism plays a role, but exclude local effects such as skin/eye irritation. A prior assumption for coverage of all endpoints is not only sometimes impractical but is also not scientifically warranted.

[26] A breakpoint refers to a point of discontinuity, change, or cessation. A chemical that identifies a turning point in a trend is called a breakpoint chemical (ECHA, 2013).

2.4 The mechanistic basis of using analogues or chemical categories

2.4.1 Principle and considerations

A category of chemicals will often show the presence and absence of a particular effect among the members of the category, based on a common functional group, physical-chemical properties, common reactivity, metabolism, and a presumed mode (and/or mechanism) of action based on a similar structure. However, a modulation of effects could appear as a result of a constant pattern in changing chemical structure or physical-chemical properties across the category. Examples can be found from the Cooperative Chemicals Assessment Programme (http://www.oecd.org/env/ehs/risk-assessment/oecdcooperativechemicalsassessmentprogramme.htm).

The read-across should be substantiated for every endpoint with data gaps depending on the regulatory programme requirements when the category and analogue approach are being applied. A chemical category approach may be suitable for more toxicological endpoints or other endpoints, since the structural changes across the category may affect changes in physical-chemical properties or other molecular descriptors or profilers that would cause changes of several toxicological properties or other endpoints of the individual category members in a coherent and consistent manner. However, it may only be possible to identify the trends and changes for some, and not all, of the endpoints of potential interest. Hence, it may not be possible to use a category approach for all relevant hazard endpoints.

When the data for a category include one or more exceptions to the effects expected from a common mode (and/or mechanism) of action, a review of the toxicological data for the category should generally be able to explain the difference in toxicity. Exceptions should not systematically be excluded from the category since the information or experimental data these provide can explain certain characteristics observed (e.g., absence of trend for a given endpoint) and may guide on the best approach to take for filling data gaps (e.g., worst case read-across vs. read-across to the closest analogue). The presence of such "outlying" effects underlines the importance of developing an understanding of the (toxic) mode (and/or mechanisms) of action within categories.

A category may be justified on more than one basis. For example, a category could be justified by both chain length and metabolic pathway (Caley *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 mode (and/or mechanisms) of action has been properly identified.

Comparable considerations also apply to the analogue approach, where, in addition to structural similarity and similar physical-chemical properties between the source chemical(s) and the target chemical, criteria such as common functional group, biochemical processes and mode (and/or mechanism) of action, or environmental fate come into play for judging the suitability of source chemical(s) . 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. An example is the category of monoethylene glycol ethers and their acetates or diethylene glycol ethers and their acetates (OECD, 2004b). Another example is alkaline properties driving the acute oral and dermal toxicity and therefore justifying the grouping of primary amines, whereas differences in metabolism (due to structural differences) between members of the category (i.e., methylamine and tert-butylamine differ from the rest of the category) lead to different patterns of effects for chronic toxicity (see C1-C13 Primary Amines) (OECD, 2011c). Meaning that the metabolites are causing the observed toxicological effect and thus the formation pattern predicts the observed toxicity.

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. For example, methyl

alcohol and ethyl alcohol exhibit differences in their acute toxicity. Differences are seen in the carcinogenic profile between butter yellow and its ethyl homologue as well as between methylcarbamate and ethylcarbamate. This may be the result of specific differences in the route of metabolism (Jäckh, 2007).

The presence of a breakpoint (e.g., structural change, change in physical-chemical properties) can indicate a change in the mode (and/or mechanism) of action, absorption, distribution, metabolism, excretion (ADME) properties, 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 mode (and/or mechanism) of action is also a factor in deciding what chemicals would not be expected to be members of a category. Variations in chemical structure can affect both toxicokinetics (e.g., 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, while halogen substituents tend to increase lipophilicity and increase toxicological activity (see example in Worth *et al*., 2007). Thiols and esters are not considered as relevant analogues for evaluation of ether activity (see example in Hart and Veith, 2007).

Several examples for selection of analogues can be found in (Wu *et al.,* 2010). The paper provides an example for interpretation on the suitability of different analogues for filling of data gaps. An example for suitable analogues is given with the alpha-terpinyl acetate and alpha-terpinyl propionate, which have one methyl group difference in the chemical structure. Alpha-terpinyl acetate and propionate are esters of the alpha-terpineol and have very similar features, reactivity, physical-chemical properties, and metabolic pathways. The likely metabolic pathways for both substances involves conversion (hydrolysis) to alpha-terpineol, which would be the key toxicological species, and the simple acids (acetic and propionic). Consequently, the metabolism will not diverge in a manner which could lead to different toxicological outcomes.

The same author described an example where dodecanoic and stearic esters of sorbitol are potential target and analogue substances. These have similar structural features, reactivity, and metabolise in a similar manner - hydrolysis of the esters to release sorbitol and their corresponding fatty acid. However, compared to the dodecanoate sorbitol, the stearate sorbitol has a longer alkyl chain that affects its physical-chemical properties. The estimated log Kow value of the stearate ester is three units higher than for the dodecanoate ester. Respectively, the water solubility is lower for the longer-chain ester. The difference in the chain length should not have a significant effect on the metabolism pathway and the mode (and/or mechanism) of action. However, the bioavailability depends on solubility and thus on the length of the alkyl chain. The two substances would be expected to have quite different absorption rates through the dermal route, but absorption during ingestion would not be expected to be so different. Thus, consideration would be needed, depending on the route of exposure and the toxicological endpoint of interest, to incorporate this added uncertainty.

The same paper also shows that a source and target chemical could have different toxicological properties as a result of significant metabolic convergence or divergence. One example of such analogues (the hypothesis being – convergency to the same or similar toxic metabolite) is a set of 4-alkylsubstituted phenols, which can be oxidised to form quinone methide derivatives and may generate similar toxic effects such as cytotoxicity and skin sensitisation, caused by the (bio)transformation[27] product. However, there are factors that could affect the formation of this product (the conditions) such as a hydrolytic enzymatic process. The usefulness of data from this type of analogue depends upon the probability that the required pre-condition will be met *in vivo*. ADME properties are an important consideration in the justification of most categories and their associated read-across. Kinetic information can help to demonstrate that substances within a category are dealt with in a similar or predictable manner by the body (i.e., the substances are metabolised in a similar manner), have

[27] (Bio)transformation is the chemical modification of a chemical within an organism.

similar absorption and excretion kinetics and similar distributions, and therefore would be expected to exhibit similar toxicological properties. Such data are also useful in substantiating any trends observed across the category. For example, toxicokinetics data can be used to confirm that as molecular weight increases, bioavailability decreases, thus reducing the potential for systemic toxicity. Having toxicokinetic information does have the potential to reduce the uncertainty in assessing the read-across and could therefore obviate the need for intensive higher-tier testing for category members. However, conducting ADME studies generally requires animals, is expensive, and the studies can be technically challenging and lengthy to run. A pragmatic solution could be to consider collecting toxicokinetic data as part of standard toxicological studies. Saghir *et al.* (2012) and Creton (2012) outlined overviews of how toxicokinetic parameters could be included in standard guideline studies from sub-acute to chronic repeated-dose toxicity studies and developmental and reproductive toxicity studies. The approach proposed considering measurements taken during the range finding of a study, therefore using a limited number of animals and avoiding the use of radiolabelled material. Obviously such an approach relies on a good hypothesis on the likely metabolic pathway in order to be able to positively identify the metabolites formed through analytical methods. Using tools such as the OECD Toolbox to identify likely metabolites and the potential rate of their formation could also help to characterise ADME properties. Further elaboration of metabolic considerations is provided in Chapter 6.

2.4.2 The concept of Adverse Outcome Pathways and the use of bioprofiling information for grouping chemicals

This section is meant to introduce how the concept of MOA/AOP and bioprofiling information can be used in forming/justifying chemical categories.

The definition of a group starts with structural similarity and allowed structural differences and then continues with investigating the hypothesis for common mode of action. The possibility to confirm a common mode (and/or mechanism) of action within a chemical category has been further investigated in the last years at OECD via the development of the concept of adverse outcome pathways (AOPs). An AOP delineates the documented, plausible and testable process by which a chemical induces molecular perturbations and the associated biological responses which describe how the molecular perturbations cause effects at the sub-cellular, cellular, tissue, organ, whole animal and (when required) population level of observation. The pathway approach is based on the concept that toxicity results from a chemical first reaching and then interacting with an initial key target (e.g., membrane, receptor) in the organism; this is defined as the primary molecular initiating event. Subsequent to this primary interaction begins a series of events that can individually be documented and tested, resulting in an adverse outcome (e.g., reproductive failure, neurotoxicity). Obviously, several pathways can result in the same adverse outcome, and each constitutes an individual AOP. An OECD workshop on using mechanistic information in forming chemical categories was held in December 2010 (OECD, 2011a). The aim of the workshop was to investigate how AOPs can be used for grouping chemicals. The report provides important definitions (e.g., molecular initiating event, toxicity pathway, mode (and/or mechanism) of action, AOP, as well as case studies on proposed AOPs.

Integrating knowledge of how chemicals interact with biological systems (i.e., the molecular initiating event) with knowledge of the responses at increasing levels of biological complexity allows for the forming of chemical categories and reliable predictions of long-term effects. Indeed chemicals can be grouped according to their ability to trigger the same molecular initiating event or following key events[28]. It is not needed to have the full pathway from initial molecular initiating event to the final adverse outcome described in a document before building a chemical category around a common mode (and/or mechanism) of action or key event. It is nevertheless necessary to establish causal links between the molecular initiating event) or key events (KE) used to group chemicals and the apical endpoint for which the data gaps are to be filled.

[28] Key events are intermediate events (i.e., events between the molecular initiating event and the apical outcome) that are toxicologically relevant to the apical outcome and experimentally quantifiable.

Results from molecular screening, *in vitro* assays, and omics methods and assays can be helpful in grouping chemicals particularly if these relate to the MIE or KE of an AOP. For example, the genetically engineered yeast-based bioreporter system for gene activation based on oestrogen receptor (ER) binding can be used to screen chemicals for the potential to be reproductive toxicants via ER-binding. In this scenario, the MIE is ER-binding, the adverse outcome is reproductive toxicity, and the gene expression is an intermediate event. Hence the justification for grouping chemicals to fill data gaps on reproductive toxicity could be strengthened by showing that all chemicals in the category have the same results in the ER-binding assay, in addition to the basic justification on similarity of category members that is otherwise required. As databases based on a particular protocol are developed, it will be possible to develop structure-activity relationships (SARs) or bioprofilers which allow for the determination of the chemical space or applicability domain for chemicals likely to be active for that biomarker based on 2D or 3D structural rules. The advantage of such a bioprofiler is to allow prediction of the likelihood of a chemical eliciting a particular AO based on the chemical being in the applicability domain of a mechanistically-relevant key intermediate event.

It is intuitive that MIEs and KEs can only be used in the context of a particular AO and a particular pathway. For example, ER-binding can be mechanistically linked to reproductive toxicity but binding to the aryl hydrocarbon receptor, another possible MIE, does not lead necessarily to reproductive toxicity nor are all reproductive toxicants ER binders (Beischlag *et al.*, 2008).

A template on building an AOP has been published (OECD, 2013b). It is structured around two main parts. The first part of the template is about the evidence supporting the AOP: concordance and consistency between the molecular initiating event , the key events up to the adverse outcome, and the biological plausibility. The second part of the template is about the chemical space of the AOP: the similarity in the molecular descriptors or profilers that are relevant for the AOP and the sub-structural fragment(s). The AOP for skin sensitization provides an illustration (OECD, 2012a).

2.5 Robustness of a chemical category and of an analogue approach

A number of factors contribute to the robustness of analogue(s) in a category or in an analogue approach. Useful considerations might include:

- Similarity in the structure and reactivity of analogue(s) / category members;
- Similarity or trend in the physical-chemical properties;
- Number of analogues/the density and distribution of the category (both in terms of the chemicals represented and the data available);
- Quality of the underlying experimental data for each of the endpoints covered;
- Presumed mechanistic basis underpinning the category or the analogue approach for a particular endpoint; and
- The quality of the data estimated by 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. The robustness of a chemical category is assessed on a case-by-case basis. To facilitate the assessment, documentation following a logical framework should be provided in every case containing the following elements as appropriate:

- The category definition: summary of common features; boundaries (e.g., in number of carbon atoms); physical-chemical properties, if applicable (e.g., boiling point); allowed variations in chemical structure; and if known, any restrictions (e.g., variations that would change the effects of a

chemical significantly compared to the other chemicals in the category). The category description determines the applicability of the category hypothesis to chemicals;

- The hypothesis for grouping chemicals together in a category: the structural and mechanistic similarities between members of the category (functional group, moiety of concern, carbon chain length, common metabolic or degradation pathways and products, common physical-chemical properties, discussion of presence or absence of effects);
- The category justification: all pre-existing experimental or other (e.g., literature) evidence that can support the hypothesis defined in the previous step. This could be similar effects in lower tier studies where these exist, availability of "bridging" studies which should be mechanistically relevant to the assessed endpoints (e.g., common *in vitro* or other type of screening studies), evidence from computational and non-computational theoretic models, common bioavailability and reactivity profiles, common mode and/or mechanism of action or AOP;
- An endpoint by endpoint justification to demonstrate that a trend or similarity exists between the analogues within the category. This is particularly relevant for the endpoint(s) with data gaps but should not be restricted to just those toxicological endpoints;
- The scope of application of the category approach: the domain for which the category approach applies and a list of endpoints where data gap filling is being proposed;
- The existence of sub-categories and the rationale for sub-categorising: for example, existence of a threshold for certain physical-chemical properties impacting solubility or bioavailability of category members and thus hazards for given endpoints; and
- The technique used to fill the data gaps: for each endpoint and chemical, the availability of experimental data and in the case where no data were available, how was the data gap filling applied (read-across from closest category member, worst case scenario trend analysis), with justification for the strategy.

The robustness of the chemical category and the analogue approach should also be assessed in view of their predictive value for filling a data gap. This assessment could be done as a tiered approach (ECHA, 2012a). The first tier addresses a series of straightforward questions on the occurrence, nature and quality of the read-across cases in a read-across proposal. It should include a check of whether the chemical identity and composition allows for application for read-across. More specific questions may include:

- Is the read-across explicitly stated and is the source substance clearly identified?
- Is the read-across used in a supporting role or it is meant to fill an endpoint on its own? The purpose of the read-across can be to replace the results of a standard experimental study entirely (i.e. stand-alone read-across), or may have supporting role. It can alternatively be part of a WoE analysis. The level of confidence in the data read-across prediction should be high if the read-across data is used as stand-alone tool for replacement of a data point;
- Is coverage of the key parameters addressed in the test that is replaced? The study with the source substance l must have adequate and reliable coverage of the key parameters as in the standard test method. Qualitative and quantitative differences in the investigated parameters should not result in an underestimation of hazard;
- Is exposure duration in the test with the target substance that is replaced of similar length as for the source substance: The exposure duration often strongly influences the types of effects observed and the sensitivity with which the effects are observed. For example, if the information requirement is for a 90-day repeated-dose toxicity study, it would normally not be possible to base the read-across on a 28-day study;

- What is the purpose of the read-across? The purpose of the intended read-across should be clearly stated (e.g., if the result of read-across and data gap filling approach is intended for classification and labelling and/or risk assessment); and
- Is the documentation provided adequate and reliable? The first tier should also analyse if more in depth assessment is required.

Scientific assessment of the robustness of the read-across could be done in a second step according to the hypothesis for grouping. Each hypothesis can be characterised by a set of specific aspects that taken together are crucial for the scientific credibility and reliability of the read-across case. These key aspects could play a central role in the assessment of the data from read-across for each hypothesis. Each of the key aspects could be related to defined possible assessment options. For example, one hypothesis could be that the chemical or biological conversion of target and source chemicals results in exposure to the same toxicants, and subsequently to the same effects. Important considerations for this scenario include whether:

- A common (bio)transformation product[29] could be formed in the first place;
- The parent molecules for both target and the source(s) could be toxic on their own;
- The target and the source(s) have (bio)transformation products, which differ and therefore can cause different effects;
- The target and the source(s) (bio)transformation products could cause similar exposure to the same tissues and organs; and
- Different metabolic pathways, which underpin the hypothesised one, cannot take place.

2.6 The interdependence between categories and (Q)SARs

The chemical category and (Q)SAR concepts are strongly connected given that the underlying basis of both is the essentially the same, i.e., toxicity is a function of chemical structure. The differences merely lie in the formality of how that relationship has been packaged. This is described in detail by NAFTA (2012). The concept of forming a chemical category and then using measured data on a few category members to estimate the missing values for the untested members is in essence an internal (Q)SAR. The reason this concept is so compatible with (Q)SAR is that this broad description of the categories concept and the historical description of (Q)SAR are the same.

A SAR is a qualitative relationship that relates a (sub)structure to the presence or absence of a property or activity of interest. The substructure may consist of adjacently bonded atoms, or an arrangement of non-bonded atoms that are collectively associated with the property or activity. SARs can be helpful in the qualitative evaluation of the analogues identified as belonging to a category.

A (Q)SAR (is a mathematical model (often a statistical correlation) relating one or more quantitative parameters derived from chemical structure to a quantitative measure of a property or activity (e.g. a (eco)toxicological endpoint). (Q)SARs are quantitative models yielding a continuous or categorical result.

Both SARs and QSARs have been implemented as “profilers” in for example the OECD QSAR (Q)SAR Toolbox, where their use can aid in the construction of chemical categories.

Similarly to (Q)SARs, a quantitative activity-activity relationship (QAAR) is a mathematical relationship, but between two biological endpoints, which can be in the same or different species. QAARs are based on the assumption that knowledge about the mode (and/or mechanism) of action, obtained for one endpoint, is

[29] A transformation product is the result product from biotransformation.

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). Examples of QAARs include relationships that allow for the prediction of daphnid toxicity from *Tetrahymena pyriformis* or acute oral toxicity from cytotoxicity measurements. While the formalisation of QAARs has not been used extensively, conceptually the use of information from one endpoint to another, where there is commonality in the rate determining step or the molecular initiating step, is well recognised. It seems that the QAAR concept has better application for environmental endpoints but for health effects its use is expected to be rather limited.

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

The categories concept creates a practical and powerful approach for describing the structural requirements of toxicity mechanisms. Chemicals may be grouped together initially using expert judgment, which is reflected by the chemicals included. Further evaluation may question the similarity of some chemicals based on measured data or evidence of anomalous behaviour or other information about the chemical attributes that suggest some chemicals may fit more than one category. Application of a (Q)SAR model(s) may be useful to help substantiate the category based on the manner in which mechanistic information has been encoded, i.e., in providing the mechanistic insight to support the interpretation of the experimental data in addition to filling data gaps themselves.

However, the errors due to the choice of a (Q)SAR model for a specific chemical can exceed the inaccuracy in the potency estimate of the (Q)SAR model. For example, in ecotoxicity studies, some phenols are polar narcotics, some are uncouplers, and others are electrophilic. (Q)SAR 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 (Q)SAR estimates are based on mechanistically valid models by aiding correct selection of the model.

Further information on the use of (internal) (Q)SARs to express trends in categories, and on the use of (external) (Q)SARs to provide additional support for trends, is given in Chapter 3, Sections 3.3 and 3.4, respectively.

For classification and labelling, internal (Q)SAR models within a category 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. 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 members to be above or below that threshold and still belong to the chemical category.

3 TECHNIQUES OR METHODS FOR DATA GAP FILLING

3.1. Introduction

In Chapter 2, both analogue and category approaches are discussed in detail and how these may be used to extend the use of measured data to similar chemicals that may not have the same level of data. In practice both approaches rely upon similar techniques to read-across data to fill identified data gaps. Consequently, this chapter on data gap filling does not differentiate between the two, but rather presents the methods that might be used to fill those data gaps whether these are via an analogue or category approach.

The OECD guidance offers general science-based advice, and it is not geared toward one specific regulatory scheme. However, examples from different regulatory jurisdictions are provided to help users understand concepts. Users of the guidance should be mindful of this and consider the aspects and requirements of the specific regulatory scheme most relevant to them.

The absence of relevant, reliable and sufficient experimental data for chemicals in a category may result in one or more data gaps that need to be filled in order to finalise the hazard and/or risk assessment. This chapter explains the following non-testing methods for filling these data gaps:

- Read-across (Section 3.2);
- Trend analysis and use of computational methods based on internal models (Section 3.3); and
- Use of computational methods based on external models (Section 3.4).

Sections 3.2.3 discuss how to develop read-across justifications for each endpoint and considerations for the validity of read-across.

In principle the above-listed non-testing techniques can be used to indicate either the presence or the absence of an effect or an estimated value (e.g., a relevant toxicity value such as a LOAEL) for an analogue or a group of substances. However, this is highly dependent on the substance under consideration, the endpoint, the level of information already available, the regulatory purpose, and the confidence that can be derived from its interpretation. Consequently, the generation of additional experimental data by strategic testing may still be required to inform the properties of category members and develop confidence in the approach considering the WoE of all of the information available.

The use of these techniques is described in more detail below. None of them are typically used in a stand-alone mode. Usually these techniques rely on building a case with varying degrees of applicability in the context of both the analogue approach and the wider category approach. Experience from current practice shows that the use of qualitative or quantitative read-across is already widely used and is a viable approach for regulatory purposes on a case by case basis. While computational approaches based on SARs, (Q)SARs, 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.

To increase confidence in the read-across approach when applied to analogues or a category, evidence must be provided to underpin the hypothesis on which the read-across is based. This can be done by adding new elements to reinforce and develop the initial hypothesis, or by providing new scientific evidence that the category parameter is behaving as expected. The most compelling evidence in support of a read-across hypothesis is information on a common mode of action of the substances and a mechanistic rationale for their common biological behaviour.

3.2. Read-across

The use of read-across is widespread across regulatory jurisdictions, particularly as a means to fill data gaps for information requirements under specific regulations. The term "read-across" is a generic and much used phrase. However, all the examples of categories and analogue approaches from the OECD HPV programme, and regulatory applications within Member Countries, make it clear that read-across can only be used on a case-by-case basis by providing a hypothesis on which the read-across is based. Adequate justification, documentation (see Section 3.2.3 for more information), and supporting data may be required for acceptance.

The principle of the read-across technique is that endpoint or test information for one chemical is used to predict the same endpoint or test for another chemical, which is considered to be similar by scientific justification. A chemical used to make an estimate can be referred to as a source chemical, and a chemical for which an endpoint is estimated can be referred to as a target chemical.

Theoretically, the technique of read-across can be applied to characterise physical-chemical properties, environmental fate, human health effects and ecotoxicity. For any of these areas, read-across may be performed in a qualitative or quantitative manner.

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

- One-to-one (one analogue used to make an estimation for a single chemical);
- Many-to-one (two or more analogues used to make an estimation for a single chemical);
- One-to-many (one analogue used to make estimations for two or more chemicals); or
- Many-to-many (two or more analogues used to make estimations for two or more chemicals).

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 "binary" or "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 (e.g., NO(A)EL, LO(A)EL). Qualitative and quantitative read-across techniques are discussed in more detail in Section 3.2.2.

Most often, structural similarity and similar properties and/or activities between the source and target chemicals are used as a basis for justifying read-across. Structural similarity provides a convenient means of identifying likely analogues. Their suitability may be evaluated by reference to one or more of the following similarity contexts:

- Common functional group(s) (e.g., aldehyde, epoxide, ester, metal ion). An example is the ethylene glycols category assessed in the Cooperative Chemicals Assessment Programme.[30]
- Common constituent or chemical classes, similar carbon range numbers. This is frequently the case with complex substances often known as UVCBs." [31]

[30] OECD Cooperative Chemicals Assessment Programme: http://www.oecd.org/env/ehs/risk-assessment/oecdcooperativechemicalsassessmentprogramme.htm
[31] OECD Cooperative Chemicals Assessment Programme: http://www.oecd.org/env/ehs/risk-assessment/oecdcooperativechemicalsassessmentprogramme.htm

- Common precursor and/or breakdown product that results via physical or biological processes (i.e., metabolic or degradation pathway similarity). This is used to examine related chemicals, such as acid/ester/salt (e.g., esters of thioglycolic acid, thioglycolic acid and its ammonium salt[32]). Additional 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 or degradation of the dye.

3.2.1 Hypothesis and evidence based approaches

The process of developing an analogue or category starts with the identification of analogues and an evaluation of their relevance with respect to one or more similarity rationales. This is the overarching hypothesis to help substantiate any proposed read-across. The evaluation of analogues will include consideration of structure, composition, physical-chemical properties, reactivity and metabolism, and mechanistic similarity as discussed further in Section 3.2.3. This document details the elements that should be considered in both hypothesis generation and in providing evidence, and gives examples wherever possible.

The hypothesis needs to be fit for a particular purpose and provide some mechanistic basis or understanding of why it is fit for that purpose. Careful evaluation of the rationale for all the various endpoints under consideration is needed to define the applicability domain of the analogue or category to ensure read-across is appropriate for the endpoints of interest. For example if the hypothesis is based on structure and mode of action, then it may be valid for certain aspects of mammalian toxicology, but not hold for environmental endpoints. If the hypothesis is based on metabolism then the read-across may only be valid for systemic mammalian endpoints or routes of administration, and local effects such as skin or eye irritation could be excluded. Interspecies differences in metabolism may need to be taken into consideration. Once a hypothesis has been developed it then needs to be examined using the available data (evidence) to see if the hypothesis is verified for the intended endpoint. This process of building a hypothesis and then testing with the available evidence can be referred to as the read-across justification for an analogue or the category justification for a group of substances.

Following hypothesis generation, evaluation of the existing, available information should provide evidence that the endpoint read-across is robust for the target substance(s). In order to develop that certainty, it may be necessary to generate further information on specific category members and for certain endpoints to provide additional data for the full category justification. The overall approach builds on a WoE to demonstrate that the read-across is robust and that the data being used are applicable to the target chemical.

The results of read-across may be used for different purposes, from screening candidates for a particular concern – which may be endpoint specific – to classification/labelling and risk assessment. Consequently the degree of certainty from the hypothesis testing stage and the WoE necessary may vary for these different needs.

The elements of read-across detailed below provide a systematic approach to building a read-across justification. The final justification needs to take into account the level of information available on a case-by-case basis and address each endpoint for which the read-across is proposed.

3.2.2 Choice of qualitative or quantitative read-across

In the OECD Cooperative Chemicals Assessment Programme where hazards are determined, a "strict" quantitative read-across is not systematically applied, and often the determination of the presence/absence of a

[32] OECD Existing Chemicals Database: www.oecd.org/env/existingchemicals/data

given hazard (e.g., skin sensitization) or the determination of "regulatory threshold" (e.g., for classification and labelling) may be sufficient. However, in the case where the category approach is used for risk assessment, the assignment of quantitative values to untested chemicals may be necessary. This section provides guidance for applying quantitative and qualitative read-across. Before deciding on the type of read-across approach that is necessary, it is important to determine why the data gap is being filled and what type of data are required. Is a specific value required or does the endpoint need to be checked against a threshold or hazard banding/cut-off (e.g., a classification banding)?

In deciding whether to use quantitative or qualitative read-across, the nature of the endpoint should also be considered. It may be expressed on a numerical or categorical scale. In most cases, a specific value is required for risk assessment, such as a NOAEC or NOAEL, an environmental half-life, or a partition coefficient. Examples of using qualitative and quantitative read-across are given in Table 1.

3.2.2.1 Challenges with read-across

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 have been generated from studies that used different test methods or species for a given 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. Because 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. For example, for genotoxicity, if 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 category members that lack data (target chemicals). 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, or there may be cases where there is sufficient information on the quantitative trend across categories that it may be desirable to segment the category with different hazard identification assigned to sub-groups within the category. In any case, the validity of the read-across and the hazard characterisation should be independent of the data requirements for the programme.

3.2.2.2 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 judgment. Several factors can be considered in making this judgment. 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 or position 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 judgment. However, it should be recognised that expert judgment may not be agreed upon by all concerned in the evaluation.

If a regulatory classification is used to express the property or effect, differences in the potency of the chemicals could be sufficient to warrant different classifications, 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 in whether source and target chemicals cause direct or indirect effects may lead to differences in classification.

3.2.2.3 Quantitative read-across

In quantitative read-across, the known value of a property for the source chemical(s) is also used to estimate a quantitative value for the same property for the target chemical, which doesn't have data for the property of interest.

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

- Using the endpoint value of a source chemical, e.g., the closest analogue in a (sub)category [33];
- Using a trend to scale the available experimental results from two or more source chemicals to the target chemical [34] (it should be noted that the trend should be statistically sound); see section 3.3
- Processing the endpoint values from two or more source chemicals (e.g., by averaging, by taking the most representative value); or
- Taking the most conservative value of the closest analogues or the most conservative value in the (sub)category[35].

Quantitative read across can also be used for complex substances/UVCBs, typically by applying data from substances with similar physical-chemical properties (e.g., substances with similar boiling ranges, carbon ranges, composition) or by applying data from key/major constituents. However, quantitative read-across for a UVCB must be done carefully and requires a sufficient knowledge about the composition of the UVCB (i.e. sufficient knowledge about the identity and properties of its constituents) and hence an understanding of the key structures that are likely to drive the behaviour and properties of UVCBs. Hence read across for UVCBs, if performed, may be more appropriate to be done in a qualitative or semi-quantitative way than attempting to fill data gaps for employing quantitative approaches (quantitative read across, trend analysis or QSAR approaches). This is further discussed in Chapter 6, section 6.5.

[33] For example, the OECD HPV Cooperative Chemicals Assessment Programme 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).

[34] For example, OECD (2006b) 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

[35] For example, the assessment within the EU Existing Substances Regulation and the OECD HPV Cooperative Chemicals Assessment 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).

Quantitative read-across can be used to fill data gaps in hazard and risk assessment. Assessment factors are often applied to toxicity values (e.g., NOECs from aquatic toxicity studies or LOAELs from repeated dose toxicity studies in rodents) to yield a dose or concentration to which humans or organisms may be exposed that is expected to be 'safe' (i.e. that does not result in adverse effects). However, the use of read-across in the development of assessment factors is outside the scope of this OECD guidance document.

Table 1 Examples of Read-Across from National and International Programmes

Target Chemical **(CAS)**	**Source Chemical(s)** **(CAS)**	**Type of Read-across**	**Purpose**	**Reference**
Phenol, (1,1-dimethylethyl)-4-methoxy (25013-15-6)	4-tert-Butylphenol (98-54-4)	Qualitative	The source chemical was used to fill a data-gap for biodegradation. It was used as part of the WoE to support the half-life in water to evaluate environmental persistence in the ecological risk assessment.	Canada 2010a
Decanedioic acid, bis(1,2,2,6,6-pentamethyl-4-piperidinyl)-ester (41556-26-7)	Decanedioic acid, 1,10-bis(2,2,6,6-tetramethyl-4-piperidinyl) ester (52829-07-9) Decanedioic acid, 1-methyl 10-(1,2,2,6,6-pentamethyl-4-piperidinyl) ester (82919-37-7) Decanedioic acid, 1,10-bis[2,2,6,6-tetramethyl-1-(octyloxy)-4-piperidinyl] ester (122586-52-1)	Quantitative	Most conservative short-term oral value from a selection of analogues was used to calculate a margin of exposure for human health risk assessment.	Canada 2010b
Total Aluminum	Various aluminum-containing compounds	Semi-quantitative	43 studies on aluminum containing compounds used to characterize level at which neurological and reproductive/developmental effects begin to be repeatedly observed in animal studies	Canada, 2010c
MAPBAP acetate (72102-55-7)	5 (Q)SAR models + gentian violet (CAS 548-62-9), malachite green (CAS 569-64-2), C.I. Basic Violet 4 (CAS 2390-59-2) and leucomalachite green (CAS 129-73-7)	Qualitative	Identify potential health effects of target chemical to inform human health risk characterization.	Canada, 2010d
Thiobis Propanoic Acid Derivatives	Propionic acid, 3,3-thiodi-, didodecyl ester (3,3-thiodipropionic acid, didodecyl ester) (CAS 123-28-4) Propionic acid, 3,3-thiodi-, dioctadecyl ester (3,3-	Qualitative	Developmental toxicity testing in DLTDP (di-lauryl-thio-di-propionate) produced no adverse results in four separate species. Also, in a 90-day repeat dose study with DLTDP, no effects on reproductive organs were observed. Since DLTDP is the smallest of the three materials it is estimated to	EPA, 2001

Target Chemical (CAS)	Source Chemical(s) (CAS)	Type of Read-across	Purpose	Reference
	thiodipropionic acid, dioctadecyl ester) (CAS 693-36-7) Propionic acid, 3,3-thiodi-, ditridecyl ester (3,3-thiodipropionic acid, ditridecyl ester) (CAS 10595-72-9)		be an appropriate conservative representative for the family.	
Trimellitate Category	Tris-2(ethylhexyl) trimellitate (ToTM) (3319-31-1)	Qualitative	Due to their higher molecular weight and bulky side chains, the remaining members of this category are expected to demonstrate a lower order of toxicity than ToTM. This is supported by a similar structural-activity relationship observed with phthalate ester compounds, i.e., the higher molecular weight phthalates (ester side chains >C7) are less active that the transitional phthalates (ester side chains C4-C6). Thus, the use of TOTM to represent the potential hazards of the other category members is a conservative position. No additional toxicity tests were proposed for this category.	EPA, 2007a
Short chain chlorinated paraffins	Alkanes, C10-13, chloro-	Qualitative	Data gap filling. The NOAEL for effects via lactation was read across from medium chain chlorinated paraffins (both within the EU Existing Substances Regulation and the OECD Cooperative Chemicals Assessment Programme).	EEC, 1993 and OECD, 2000a
P-t-butylphenol CAS 98-54-4	p-t-pentylphenol CAS 80-46-6	Qualitative	To flag a concern for further testing. Data from p-t-pentylphenol were used to request further testing on endocrine disruption in fish.	EEC, 1993 and Tsakovska and Worth, 2007

3.2.3 Elements for a read-across justification

In developing an approach for data gap filling using either the analogue or chemical category, a number of general elements should be considered and discussed to demonstrate the relevance of the analogues such that the subsequent endpoint read-across is scientifically justifiable. No read-across justification is ever identical because of the nature of substances and chemical classes, and the fact that these may be grouped by a variety of means, from structural similarity, common production, common toxicological or environmental properties, through to uses and applications.

The following are general elements that can be considered when an analogue/category approach is being developed. A full read-across justification would consider a number of these elements, for a given endpoint. Endpoint-specific elements are discussed in the appendix. The elements presented here do not constitute an exhaustive list; each may serve to inform the underlying analogue/rationale.

- Chemical identity and composition
  - Chemical structure
  - Composition
  - Impurities
  - Functional groups
- Physical-chemical properties and other molecular descriptors [36]
- Kinetics: Absorption, distribution, metabolism and excretion
- Mode/Mechanism of action or adverse outcome pathways (MOA/AOP)
- Chemical / biological interaction
- Responses found in alternative assays
- Information obtained from other endpoints/species/routes

These elements are discussed in the following sections. Examples on how these have been used in previous category assessments are provided whenever possible.

It should be noted that the order of this list does not necessarily imply a hierarchical approach, even though some hierarchy of considerations could be applied. For example, although structural similarity is very often a starting point for read-across, it may not always provide the best scientifically supportable basis for determining analogue relevance or group membership. However, the similarity in chemical structure is the first requirement to consider, assuming that the composition is analysed and impurities or other constituents are not expected to change the toxicity profile. Detailed explanation for how chemical similarity determines the toxicity or the lack of it is particularly necessary for complex toxicity endpoints (e.g., cancer, reproductive toxicity) that are linked to multiple mechanisms. To the extent possible, the read-across should be done by linking structural moiety(ies) in the source and target chemicals and their tautomers/isomers, metabolites, degradation products and derivatives to specific common mechanisms.

3.2.3.1 Chemical identity and composition

3.2.3.1.1 Chemical structure

The chemical structure of substances usually provides the initial rationale and impetus for developing an analogue or category approach. However, similarity of chemical structure is hardly ever a complete justification for a category. Consequently, some categories that are constructed for purposes of assessment are likely to be composed of a subset of all the potential structures that could be envisaged.

On a case-by-case basis, and when available and not proprietary, production process chemistry may inform the understanding of common structural elements within a category; similarly, variation in manufacturing processes may result in differences in chemical composition. This applies especially for UVCBs. For discrete chemicals, composition and impurities are supposed to be known and defined. Process

[36] Molecular descriptors are numerical quantities describing the chemical structure and can be divided into two categories: Classical physical chemical properties (e.g. solubility, log P, molar refractivity, dipole moment, polarizability), and theoretical molecular descriptors derived from a symbolic representation of the molecule (0D, 1D, 2D and 3D, e.g. counts of structural fragments and atoms, connectivity indices and other graph invariants, HOMO/LUMO, surface area, volume).

chemistry may also provide useful information on composition, purity and physical-chemical boundaries, especially for process chemicals with a less-defined chemical structure.

The chemical structure(s) need to be described in sufficient detail to convey an understanding of the elements that will affect the properties of the category members and set boundaries for the category. The exact nature of these will be dependent upon the category and its chemistry, but will include one or more of the following elements:

- Overall structural trend and/or structural similarity
- Functional group(s)/Moieties
- Carbon chain length
- Linearity, branching
- Degree and position of branching
- Cyclics and aromatics
- Isomerisation
- Salts and their relationship to a source chemical / parent substance
- Purity(ies)/Impurities

3.2.3.1.2 Composition

For some categories it is not possible to define the structures as detailed in the previous Section. This is because the categories do not contain single constituents, but instead multi-structural substances. If the exact structure is variable or not known, these substances are typically referred to as UVCB (unknown or variable composition, complex reaction products or biological materials) substances.[37] In the EU, substances with more than one constituent but known composition are referred to as multi-constituent substances (ECHA, 2012b).

Even though UVCBs vary in composition, it may be possible to define how UVCBs of a similar nature will act by demonstrating the similarity of composition between the source and the target chemicals. The challenge is that unlike categories of substances with defined chemical structures, UVCB categories are often not progressive in their relationships between chemical structures and can be visualised as more of a "cloud" or "sphere" of substances with overlapping chemistries and properties. Consequently, a more expanded definition

[37] UVCB substances (unknown or variable composition, complex reaction products or biological material). 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. There are many different types of complex substances, although generally these all have the following characteristics in common.

- These 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.
- These are not intentional mixtures of chemicals.
- Many are of natural origin (e.g., crude oil, coal, plant extracts) and cannot be separated into their constituent chemical species.
- The concept of "impurities" typically does not apply to complex substances in some OECD Member States.
- These are produced according to a performance specification related to their physical-chemical properties.

of the boundaries of their chemistries compared with the boundary definition of categories of discrete chemicals is a useful way to define a UVCB category. See for more guidance in section 6.6.

It is not the purpose of this section to detail all the varieties of UVCBs and multi-constituent substances and how these might be addressed in detail. However, in building a category justification for read-across, it is necessary to be able to define the boundaries of the chemistry within the UVCBs so that some analysis can be made of the validity of the proposed read-across.

Two generic approaches can be envisaged to assist in this process:

One approach is to conduct an assessment of the composition and trends of existing data among the category members, in order to understand the boundaries of the category. Within the intended UVCB or multi-constituent category, various components of the complex substances can be further assessed for any trends and the limits of those blocks defined. Such an analysis will yield whether target and source chemicals are indeed related and which ones will provide a valid read-across between one another. Which chemical blocks are used to help identify the category will be highly dependent on the nature of the substances involved. For some it will include specific chemical properties, and for others it will rely on blocks of similar structures.

A second approach is based on well-known hazards of constituents. For example, if the structures in the UVCB are known, then it may be possible to define which ones are of high hazard, and therefore would drive a hazard assessment. For example, if a petrochemical stream contains benzene or polycyclic aromatic hydrocarbons, recognised human carcinogens, their content in the UVCB would likely drive any human hazard assessment and there may be limited value in considering any other information. Furthermore, if there are reliable data on the petrochemical stream as a whole, that data should be considered in a WoE. Only relying on constituent data would lead to less informed hazard characterizations.

3.2.3.1.3 Examples of categories and structural relationships

A number of UVCB categories have been assessed within the OECD programme, three of which are described in Table 2.

Table 2. Examples of categories based on component analysis

Category	Constituents	Supporting information	Reference
C10-C13 Aromatics Hydrocarbon Solvents Category	Defined aromatic (single and double) hydrocarbon members with specific constituents / component profiles or composition Defined boiling point range (manufacturing process defines product) Carbon number range of category members identifies at minimum approximately 80% of the chemical constituents in the substance	Carbon number ranges for all members Boiling point ranges for all members Maximum 10% naphthalene Other purity criteria	SIDS Initial Assessment Profile, OECD, 2012b

C14+ Aliphatics Hydrocarbon Solvents (<2% aromatics)	Defined for members with specific constituents / component profiles or composition Contains multi-constituent substances (UVCBs) that have a variable composition due to their chemistries and method of manufacturing	Compositional analysis by: Carbon number Boiling point range Maximum levels of aromatics , sulphur and nitro compounds	SIDS Initial Assessment Profile, OECD 2011d
Anthracene oils	Applicability – environmental endpoints only. Major components – various anthracene ring structures Minor constituents – 3 to 4-fused aromatic sulphur-, nitrogen- or oxygen-heterocycles	Main components and concentration ranges for all category members	Initial Targeted Assessment Profile (Environment), OECD 2009c

Note: The selection of the examples in Table 2 was on an empirical basis and is intended to be illustrative of the types of chemistries and structures that have used the category approach for read-across. In the development of the highlighted categories other elements were also used in the justification for any read-across. The full category justification given in the reference should be consulted.

3.2.3.1.4 Characterising composition

When developing a category on the initial premise of composition for multiconstituent or UVCB category members, the compositional elements need to be stated in a clear and unambiguous manner together with the relationship between the various category members. The clearest way to do this is to build a table of the category members that indicates the elements that change and those that stay constant within the category and provide some indication of structure.

Table 3. Example for characterising category composition

Category member	Compositional Element 1 % Typical Concentration Range	Compositional Element 2 % Typical Concentration Range	Compositional Element 3 % Typical Concentration Range	Other common features from identification, production, or materials
1	10 – 20	60-80	10 – 20	Example: Produced by the reaction of R-x with y under condition of…
2	15-30	85-90	-	
3	60-80	10-20	0-10	

Common features may include such aspects as

- Ranges of a particular composition (e.g., typical ranges)
- Production orientated information
  - boiling point ranges
  - acid numbers

- carbon chain length

Just as when reporting structural elements, the objective is to build an overall picture of the domain of the proposed category and identify or quantify to the best individual members, defining the relationships between its members and setting the boundaries of its chemical properties. In order to make the case for the category definition on compositional grounds, relevant supporting analytical data and physical-chemical properties may be required to demonstrate how compositional properties change and bound the category.

3.2.3.1.5 Impurities

Purity of the substances in a category is related to the structure and the manufacturing processes that produce the compounds. When assessing any substance, it is necessary to understand whether impurities may affect that assessment. This is equally important within a category as data generated on chemicals with impurities that affect the intrinsic hazard of the sample tested could be inadequate source data for the rest of the category. Conversely, it is not possible to predict the intrinsic hazard of a target substance if it has a significant impurity of unknown hazards, unless the source and target substance have the same, significant impurity.

Consequently, when building a category one must consider the issue of impurities and decide whether it is necessary to set limits on purity levels of the chemical(s) to ensure validity for any future read-across. A number of regulatory programmes of OECD member countries consider information on levels of impurities and how these affect intrinsic properties. For example, the EU Classification, Labelling and Packaging Regulation EC (CLP-Regulation (EC) No 1272/2008) includes provisions for the differential classification of substances based upon the presence of certain impurities/constituents with specific concentration limits, although this approach in setting purity limits is not consistent globally. The use of constituent data to classify complex substances should also consider if data/information are available for the complex substance as a whole to assess hazards; only relying on constituent data would lead to less informed hazard characterization (IPIECA, 2010).

3.2.3.1.6 Functional Groups

The common functional group(s) (e.g., aldehyde, ester, epoxide, metal ion) within a category is(are) likely to be one of the critical elements that defines a category, and chemical structure also informs the likely physical-chemical nature and biological reactivity of the category members. Some functional groups are "chemical alerts" that should be explicitly identified. Such groups may require more detailed examination and analysis through the use of modelling tools.

An attempt should be made to explain how variations in the category member structures will impact the particular functionality that is identified. For example, a complex series of esters may increase significantly in molecular weight, leading to structure folding and thus, decreasing availability of the ester moieties for hydrolysis or enzymatic activity. Such a situation would imply that the intrinsic hazard could be significantly different for smaller versus larger category members. This difference could require a different approach and explanation, and provide a basis for the boundaries of the category or a subcategory.

Some functional groups or moieties may act as "interfering groups" blocking activity or changing the pattern of a biological response. For example, a bulky ester may not be hydrolysed as it will not interact with the esterase due to its size. Alpha-beta-unsaturated alcohols can be metabolically activated to form the corresponding alpha-beta unsaturated aldehydes or ketones. The latter may undergo Michael-type addition reaction due to the activated double or triple bond. In such cases the rationale for any expected change in activity needs to be documented and the effect on predicted properties explained.

3.2.3.1.7 Examples of categories and structural relationships

The majority of categories that have been examined by OECD and by member country regulatory programmes were initially conceived from structural considerations.

A selection of examples of the structural basis of categories from more than 20 years of the Cooperative Chemicals Assessment Programme can be found at www.oecd.org/env/existingchemicals/data and a number are listed in Table 4 with links to the published documentation that gives the associated detailed rationale. It should be stressed that in all of these cases, the structural consideration was a starting point and not the entire justification. Further evidence to demonstrate why the read-across is valid between category members is always necessary. It should be noted that this table does not include an exhaustive list of categories conceived from structural considerations.

Table 4 - Examples of structures and functional groups in selected categories from the OECD HPV Programme

Category	Structural Relationship between category members	Functional groups	Number of substances in the category	Major justification for category	Reference*
Alkyl chlorosilanes	Chlorosilanes, mono, di and tri	R-Si-Cl	4	Chemicals grouped based on similar molecular structure, high reactivity, physical-chemical and toxicological properties.	SIDS Initial Assessment Profile, OECD 2010b
Alkyl Sulfates, Alkane Sulfonates and alpha Olefin Sulfonates	Alkyl sulfates with a predominantly linear alkyl chain length of C8-C18, C8-C18 alkane sulfonates, and alpha-olefin sulfonates with linear aliphatic chains of typically C14-C18.	R-SO_3-cation	57 (43 alkyl sulfates, 6 alkane sulfonates, 8 alpha olefin sulfonates)	Presence of predominantly linear aliphatic hydrocarbon chain with a polar sulfate or sulfonate group, neutralized with a counter ion (i.e., Na+, K+, NH4+, or an alkanolamine cation) is most important common structural feature. Close structural similarities result in physical-chemical properties and environmental fate characteristic that follow regular patterns. Common physical and/or biological pathways result in structurally similar breakdown products and, with the surfactant properties, are responsible for similar environmental and very similar hazard profiles. Structural similarities result in same mode of ecotoxicological action. Varying length of the alkyl chain is most important parameter influencing ecotoxicity within each subcategory.	SIDS Initial Assessment Profile, OECD 2007b
Alkylamidopropyl betaines	Amphoteric surfactants	quaternary ammonium ion, carboxylic structure, amide bond	3	Category members are amphoteric surfactants containing a quarternary ammonium ion, a carboxylic structure, and an amide bond and are all manufactured from oils, usually coconut oil containing mixtures of C8 to C18 fatty acids and marketed as aqueous solutions (20 - 40 %). Structural and functional similarities and comparable physical-chemical properties of cocamidopropyl betaine inner salts and sodium salts suggest a similar ecotoxicological and toxicological profile. Values for physical-chemical endpoints for lauramidopropyl betaine are similar or within the range of values for cocamidopropyl betaines, supported by accepted (Q)SARs, therefore similar ecotoxicological properties are assumed. All available physical-chemical and environmental fate data are similar for lauramidopropyl betaine and cocamidopropyl betaine. MOA for aquatic toxicity should be the same because only the alkyl chain length differs for the chemicals in the mixture.	SIDS Initial Assessment Profile, OECD 2006c

Category	Structural Relationship between category members	Functional groups	Number of substances in the category	Major justification for category	Reference*
Alpha olefins	Even-numbered, unbranched aliphatic chain(C6 – C14) with no other functional groups	alpha- Olefin CxH2x	5	Category members are olefins bearing a single medium-length (C6 – C14), even-numbered, unbranched aliphatic chain with no other functional groups. There is an increasing or decreasing trend or pattern from the shortest category member (C6) to the longest category member (C14) for various physical- properties and ecotoxicity but there appears to be no difference across category members for biodegradation and health endpoints. Melting point, vapour pressure, and water solubility decrease with increasing chain length while boiling point and octanol:water partition coefficients increase with increasing chain length. Measured and predicted acute aquatic toxicity data indicate that 1-hexene, 1-octene, and 1-decene exhibit acute effects to aquatic organisms at levels at or below their water solubility, but 1-dodecene and 1-tetradecene are not likely to be acutely toxic. 1-hexene may be less toxic than the rest of the category members and 1-octene, 1-decene, and 1-dodecene are expected to be similarly toxic. Modelling could not predict the chronic aquatic toxicity of 1-tetradecene. No apparent difference regarding biodegradability. Data indicate no differences among the five category members for acute toxicity, repeat dose toxicity, genotoxicity and reproductive/developmental toxicity.	SIDS Initial Assessment Report, OECD 2001a
Aluminum alkoxides	Category members comprised of inorganic component and linear alcohol component		2	Category members have low toxicity to human health. Aquatic toxicity varies depending on the carbon chain length. All alcohols biodegrade and are not persistent.	EPA, 2008
Amine oxides	Alkyl hydrophobic substituent of different chain lengths with polar "head"	Amine oxide $R_3N \rightarrow O$	15	Category members have similar structures and functions. Substances are surfactants with a polar "head" (the amine oxide) and a relatively inert, hydrophobic "tail" (the long alkyl substituent). Structural variations are three-fold: 1) the nature of the second and third substituents on the amine are either methyl groups or hydroxyethyl groups; 2) the long alkyl chain ranges in length from 8 to 20 carbons; and 3) the long alkyl chain may contain one or two double bonds. Alkyl chain lengths range from 8 to 20 with 12 and 14 being predominant. Average chain lengths for the mixtures are 12.9 to 13.5, with the exception of one tallow-derived compound. Presence of methyl- vs. hydroxyethyl-substituents affects the basicity of the nitrogen only marginally, and the hydroxyethyl group lends more bulk to the hydrophilic head-group of the surfactant. Length of the longest alkyl substituent does not alter the chemical reactivity of the molecule, but does affect its physical properties. Influence of unsaturation in the alkyl expected to make the molecule prone to reactions as typical for unsaturated fatty alkyl chains.	SIDS Initial Assessment Report, OECD 2006d

Category	Structural Relationship between category members	Functional groups	Number of substances in the category	Major justification for category	Reference*
Benzyl derivatives	10 substances in category all contain benzene ring bonded directly to oxygenated functional group (aldehyde or ester) hydrolyzed and/or oxidized to benzoic acid derivative	Oxygenated functional group	10	Category members are structurally similar; but substituents and functional groups are different enough that the aldehydes, phenols and esters could each exhibit different toxicities and sensitivities. Chemicals divided into subcategories. Based on the differences in the substituents.	EPA, 2010 http://www.epa.gov/hpv /
Butenes	Isomeric differentiated C4 hydrocarbon isomers – same chemical formula and one double bond between two carbon atoms		6	Category members similar from process and toxicological perspectives. Members share somewhat similar physical-chemical properties, suggesting similar environmental fates and kinetic properties. No specific target organ was identified and no (or minimal) changes in body weight were found at the highest dose only for all the chemicals.	SIDS Initial Assessment Profile, OECD 2004c
C2-C4 Aliphatic Thiols	Straight or branched aliphatic carbon chain	Sulfhydryl functional group R-S-H	4	Category members contain a sulfhydryl functional group with a straight or branched aliphatic carbon chain. All are soluble in water and have comparable melting points, initial boiling points, vapour pressures, and low and objectionable odor thresholds. Water solubility and narrow range of octanol-water partition coefficients for the three linear C2-C4 Aliphatic Thiols indicate similar environmental fate and are not expected to bioaccumulate in aquatic organisms. Ecotoxicity is similar for the three linear members. Toxicology data show that the C2-C4 Aliphatic Thiols also have a similar order of toxicity.	SIDS Profile, OECD 2010a
Chloroformates	Alkyl chain with chloroformate group	-(–O(C=O)-Cl)	7	Chemicals grouped based on similar structure (i.e., the chloroformate group), high reactivity of the chloroformate group, and toxicological and environmental effects. Category justification based not only on similar structure but also on similar mechanism of action that results in similar human health and environmental effects.	SIDS Initial Assessment Profile, OECD 2010c
Cinnamyl derivatives	4 substances in category contain either 3-phenyl-2-propenal or 3-phenylpropanal backbone		4	Substances grouped based on close structural relationships and resulting similarities in physical-chemical and toxicological properties. Common structural features among members of this chemical category are that these contain either a 3-phenyl-2-propenal or 3phenylpropanal backbone.	EPA, 2000 http://www.epa.gov/hpv

Category	Structural Relationship between category members	Functional groups	Number of substances in the category	Major justification for category	Reference*
Dicarboxylic acid	Category composed of linear alkanes with common functional group (carboxylic acid) at each end of alkane chain		3	Category members have a common functional group, i.e., carboxylic acid, at end of alkane chain. Materials change by increase in carbon number from addition of CH_2 in alkane chain between carboxylic groups. Terminal carboxylic acids and limited chain length yield similar structure relationships. Members also share similar physical-chemical, environmental fate, ecotoxicological, and mammalian toxicological properties.	EPA, 2001 http://www.epa.gov/hpv
Ethylene glycols	Two terminal hydroxyl groups; the only variation is in the number of oxyacetylene units	glycols HO(CH2 CH2 O)n H, where n = 1-5	5	All category members have two terminal hydroxyl groups and differ from each other in the number of oxyethylene units. Members are therefore closely related in structure and have physical-chemical properties that differ as expected resulting from increasing molecular weight and consistent functionality of hydroxyl moiety on each end of molecule. Hazard and dose response profile expected to change consistently, as confirmed by data and modeling.	SIDS Initial Assessment Profile, OECD 2004b
Fuel oils	8 ethylene industry streams consisting predominantly of same higher-boiling hydrocarbons (mostly cyclic olefins and aromatics)		8 streams	Members are complex substances, containing variable amounts of alkanes, cycloalkanes, aromatics, olefins, asphaltenes, and hetero-molecules containing sulfur, oxygen, nitrogen and organo-metals. Typically defined by process history, physical properties, and product use specifications. Streams that have undergone similar processing have similar physical/chemical/biologic properties and environmental fate and transport characteristics. Refinery streams within the heavy fuels category can therefore be grouped into seven subcategories based on their process histories.	EPA, 2004
Long Chain Alcohols (C6-22 primary aliphatic alcohols)	Same basic structure CH3(CH2)nCH2OH with variations with alkyl or methyl branching and some unsaturation in some of the 30 category members	R-OH	30	Category members share the same structural features, similar metabolic pathways, common mode of ecotoxicological action, and common levels and mode of human health related effects.	SIDS Initial Assessment Profile, OECD 2006b
Nitrates	Nitrate salts	Nitrate NO_3^-	7	Members are all inorganic salts which are solid under ambient conditions (except UAN, which is a solution). Considered part of same category based on similar environmental fate, ecotoxicological and toxicological properties.	SIDS Initial Assessment Profile, OECD 2007c

Category	Structural Relationship between category members	Functional groups	Number of substances in the category	Major justification for category	Reference*
Substituted p-phenylenediamines	Category members phenylenediamines with various substituent groups always in para position of aromatic ring. Substituent groups may be all alkyl, all aryl, or mixed alkyl/aryl.		7	Category members share similar structures, physical-chemical, and toxicological (including ecotoxicological) properties, and are not readily biodegradable.	EPA, 2001 http://www.epa.gov/hpv
Sulfosuccinates	Category members have succinic ester backbone in which carbon alpha to one of the carboxyl functions has sodiumsulfo group in place of hydrogen atom		3	Category members have similar structures, follow a pattern regarding physical-chemical properties and ecotoxicological endpoints, and share similar toxicological properties.	EPA, 2001 http://www.epa.gov/hpv
Xylenes	Dimethyl benzene isomers	Bz–Me	4	Ortho- meta- and para-xylene are chemical isomers and the only difference is the position of the methyl group on the benzene ring. Mixed xylene is a mixture of the three isomers and in addition, typically contains 15-20% ethylbenzene. Category members share similar physical-chemical properties with the exception of the higher melting point of p-xylene. Toxicity of three individual isomers and mixed xylene is also similar.	SIDS Initial Assessment Profile, OECD 2003
C1 -13 Primary Amines	Alkyl amines – Increasing alkyl chain and branching	primary amino-group RNH2,	11	Category members are structurally similar with trends physical-chemical properties and ecotoxicity and similar toxicological properties.	SIDS Initial Assessment Profile, OECD 2011c
High molecular weight phthalate esters	Phthalic acid esters with carbon backbone R =>7	Ph-C-CO-O-R	7	Category consists of esters with alkyl carbon backbone with 7 carbon (C) atoms or greater. Category members contain linear and/or branched diheptyl, dioctyl, dinonyl,didecyl, diundecyl, didodecyl, and/or ditridecyl phthalate esters. Members also generally similar with respect to select physical-chemical properties or display an expected trend. Members also similar regarding biological activity, i.e., these demonstrate few biological effects.	SIDS Initial Assessment Profile, OECD 2004d

*References – *Category SIDS Initial Assessment Profile OECD – Year and Link within the OECD Existing Chemicals Database http://webnet.oecd.org/Hpv/ui/Default.aspx*

Note: The selection of the examples in Table 2 was on an empirical basis from different regulatory and international programmes and is intended to be illustrative of the types of chemistries and structures that have used the category approach for read-across. In the development of the highlighted categories other elements were also used in the justification for any read-across. The full category justification given in the reference should be consulted.

3.2.3.2 Physical-chemical properties

Physical-chemical parameters are one critical determinant to the environmental and health properties of a substance affecting bioavailability, environmental fate, and thus the (eco)toxicity of a chemical. Consequently the similarity (or logical trend) among the physical-chemical properties of category members is an important element in building a read-across approach. If the source and the target chemical share similar properties, it might be hypothesised that there is a similarity of uptake and distribution in tissues of living organisms. Nevertheless, chemicals with equal physical-chemical properties may still have different interactions with enzymes that could result in different metabolism and thereby distribution and elimination.

The most used properties in this regard are shown in Table 5 below.

Table 5. Physical-chemical properties important for hazard assessment

Property	*Related to*
logKow Log Koa,	Adsorption, estimation of bioconcentration in gill respiring animals, aquatic toxicity, mammalian absorption (oral and dermal) Estimation of potential for bioaccumulation of non metabolisablemetabolizable substances in air breathing animals
Water solubility	Adsorption, Henry's law constant, aquatic toxicity, hydrolysis
Molecular weight Molecular dimensions	Bioavailability, absorption or bioaccumulation, steric hindrance i.e. 3 D structural characteristics such as Dmax and molecular length (distribution or probability)
Vapour pressure Henry´s Law constant	Volatility with respect to choice of test conditions, inhalation Distribution coefficient between air and water, potential for exposure from water based formulation and hence relevant for considering inhalation route of exposure.
Acid dissociation constant (pKa)	Degree of ionization, relationship to irritation and corrosion, hydrolysis of ionisable substances (see sections 3.3.6), potential for uptake (including bioconcentration and accumulation), and sorption to soil (e.g. clay)
Log D (calculated)	Lipophilicity, solubility, absorption, membrane penetration, plasma protein binding, distribution

For the environmental compartment, the type of supporting information that is appropriate to report will depend on the environmental endpoint intended to be read-across. However, basic physical-chemical properties that determine environmental distribution and fate (e.g., MW, water solubility, partition coefficients such as log K_{ow}) will generally be useful. Particle size and structure are also 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 would provide useful supporting information. For example, experience within the EU "New chemicals" programme suggested 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 in the acute *Daphnia* test, then further testing for other endpoints may be appropriate (Hanway and Evans, 2000).

While basic physical-chemical properties (such as those listed in Table 5) are important factors in determining the boundaries of a given category, there may be practical problems for certain classes of chemicals such as gases, surfactants, perfluorinated substances and especially for UVCBs. The use of read-across from source to target chemicals will be required. (Q)SAR predictions may also be helpful if such a prediction provides a sufficient level of confidence to understanding potential trends within the category. It should be noted that these trends sometimes may not appear linear.

A number of categories within the EPA High Production Volume Challenge Program have used the similarity of physical-chemical parameters as part of the rationale for the applicability of the category approach. One such example is that of Dicarboxylic acids in the EPA HPV program (EPA, 2001). Category members composed of linear alkanes with common function group (carboxylic acid) at end of an alkane chain. The physical- chemical properties followed general trends; with boiling point, and logKow increasing with carbon number and vapour pressure and water solubility decreasing with increasing carbon number. These trends in the physical-chemical properties had a relationship to the some of the biological properties of the category members, i.e. acute mammalian toxicity and severity of ocular irritation decreases with increasing carbon number.

The similarity in physical-chemical properties is very closely related to the trend analysis described in the preceding section for such endpoints and can be reported in a similar manner. However, the source of the information needs to be clear and whether the value is measured or calculated. Two common sources of models to predict physical-chemical properties are the OECD (Q)SAR Toolbox[38] or EPISuite™ Epiwin[39]. The physical-chemical properties are usually reported in a table. A plot of a trend is usually very helpful.

3.2.3.3 Absorption, distribution, metabolism and excretion

For the members of a given category, if it can be demonstrated that there is similarity in absorption, distribution (presence in circulatory system, target organs), metabolism (rate of metabolism and identification of metabolites), and excretion (collectively known as ADME), then it may enhance the credibility of the read-across between a source chemical and a target chemical for defined routes of exposure and endpoints.

[38] OECD QSAR Toolbox: http://www.oecd.org/env/ehs/risk-assessment/theoecdqsartoolbox.htm
[39] EPA EPISuite US EPA. 2012. Estimation Programs Interface Suite™ for Microsoft® Windows, v 4.11. United States Environmental Protection Agency, Washington, DC, USA.: http://www.epa.gov/opptintr/exposure/pubs/episuite.htm

This mechanistic basis of using analogues or chemical categories requires evidence and experimental data to underpin it. Application of mechanistic techniques is introduced with examples in Chapter 2, Section 2.4.

The following elements should be considered to support read-across when using ADME data:

- The applicability of the available data from the source to the target substance;
- The need for supporting information on ADME on both the source and the target substance to ensure that hypothesis is valid;
- The influence of variables in the category that are also changing, for example molecular weight and logKow, and how these may influence read-across;
- The applicability of the ADME data to the appropriate route of exposure(s); and
- The applicability of the ADME data for specific endpoints.
- The applicability of the species in which the AMDE data is obtained.

For a metabolic category, it is hypothesised that a substance A is metabolised to a series of other substances and that therefore the hazard data from the substance A can be used to identify the hazard of the metabolites and vice versa. It is likely that such a category is limited to a set of primary metabolites, because primary metabolites are likely to act more similarly to the parent compounds than are more distant (e.g., secondary) metabolites. In addition, Phase II metabolites (conjugates) are generally accepted as not toxic and readily excreted. However, care should also be taken when reading across among the parent and primary metabolites. For example, it may not be useful to use data derived only from a single primary metabolite to represent toxicity of the parent compound (and/or other metabolites) because such data would not reflect the additional activity/toxicity associated with other metabolites of the parent chemical, or the parent chemicals themselves. Kinetics should consider exposure duration of the parent and or metabolite(s) which may be critical for the possibility of employing the read across approach. When knowledge about an endpoint of concern for one metabolite is available and not for the parent compound, it is essential that the exposure time to the parent compound compared with that of the metabolite is very short in order to use the data available for the metabolite to read-across to the parent compound. It should be plausible that such a short exposure to the parent compound would not lead to effects for the endpoint in question.

If data for a single metabolite are to be used, it is necessary to build the argument, not only on the metabolism data but on other criteria as well, such as knowledge that the metabolite is the only metabolite associated with toxicity or that other metabolites are likely to contribute only slightly to the overall toxicity. In such cases, additional information may also need to be considered, such as chain length and common functional groups. The chain length may result in lack of metabolic transformation (for this reason the octylchlorosilane was not included in the OECD Alkyl Chlorosilanes category, 2010), or change the metabolism (for this reason the tributylamine was not included in the OECD Tertiary amines category, 2012c). Addition of similar substances with analogous data would also broaden the general confidence in the approach. For example, a similar category of linear alcohols and their acetates or propionates could help inform the category in question (ECETOC, 2012) (see Table 6). Further information on metabolic categories are discussed in Chapter 6.

Table 6. Examples of categories with similarities in ADME/bioavailability

Category	ADME	Common structural features	Reference
Cyclohexyl Derivatives Category	4-tert-butylcyclohexyl acetate will undergo hydrolysis to yield 4-tert-butylcyclohexanol. Subsequently 4-tert-butylcyclohexanol is conjugated with glucuronic acid to yield the corresponding glucuronide that is excreted mainly in the urine.	This category consists of 2 substances, 4-tert-butylcyclohexanol and its corresponding acetate ester, 4-tert-butylcyclohexyl acetate.	EPA (2007b)
Pyridine and Pyridine Derivatives	Both piperidine and pyridine are readily absorbed through the gastrointestinal tract, skin and lungs, and eliminated primarily via the urine. Although these do not have a common metabolite, both chemicals have been shown to undergo metabolism via C- oxidation and N-oxidation, and N-methylation has been shown to be a metabolic route for pyridine. Therefore, piperidine would be expected to be metabolized and eliminated in a similar manner and rate as pyridine.	All members of the Pyridine and Pyridine Derivatives Category are structurally-related derivatives of pyridine in that these are based on the pyridine unsaturated ring structure. Piperidine (CAS RN 110-89-4) is simply the saturated ring structure derivative of pyridine.	EPA (2004a)
Terpenoid Primary Alcohols and Related Esters	Geranyl acetate is rapidly hydrolysed. The alcohols geraniol, nerol, and citronellol are efficiently detoxicated by two principal pathways. In one route, the alcohols are successively oxidized to the corresponding aldehydes and carboxylic acids, the latter of which are selectively hydrated or reduced. In a second route, the aldehydes undergo reduction to the corresponding alcohols that are substrates 11 for omega-oxidation to eventually yield diacids and their reduced or hydrated analogues. Polar metabolites formed via these two pathways will be efficiently excreted primarily in the urine as the glucuronic acid conjugates.	Citronellol, geraniol, and nerol are close structural relatives. Nerol and geraniol are cis/trans isomers of 3,7-dimethyl-2,6-octadien-1-ol and citronellol is the dihydro analogue of geraniol (3,7-dimethyl-6-octen-1-ol).	EPA (2004b)
Sulfosuccinates	It is likely that these esters will be metabolized in rodents by esterases. Compounds formed from deesterification will be similar for all three molecules, with the exception of the alcohol moiety. Whereas de-esterification of sodium diethylhexyl sulfosuccinate gives rise to 2-ethylhexanol, similar metabolism of sodium dicyclohexyl sulfosuccinate leads to the formation of cyclohexanol. Likewise, metabolism of sodium 1,3-dimethylbutyl sulfosuccinate leads to methyl isobutyl carbinol.	The general structure for the category is defined as dialkyl sodium sulfosuccinate or dicycloalkyl sodium sulfosuccinate. This describes a molecule witha succinic ester backbone, in which a carbon alpha to one of the carboxyl functions has a sodiumsulfo group in place of a hydrogen atom.	EPA (2002a)

Category	ADME	Common structural features	Reference
Phosphoric Acid Derivatives	Metabolism studies conducted on the tributyl phosphate indicate that dealkylation to form the alkyl alcohol is the primary route of metabolism. The phosphoric acid tri-esters are rapidly metabolized to di-esters with mono-diesters also being produced. Studies of tributyl phosphate show that 40-64% of the parent compound is metabolized to dibutyl dihydrogen phosphate and that 1 1 -2 1 % is metabolized to the monobutyl species. Therefore, tris(2-ethylhexyl) phosphate is expected to be metabolized to bis(2-ethylhexyl) phosphate (CAS# 298-07-7) and mono(2-ethylhexyl) phosphate (CAS# 1070-03-7). Based on the evidence for dealkylation as the primary metabolic pathway, 2-ethylhexanol is the expected metabolite of tris(2-ethylhexyl) phosphate (CAS# 78-42-2) and 2-ethylhexyl phosphate (CAS# 12645-31-7). Triisobutyl phosphate is expected to be metabolized similarly as tributyl phosphate, with methoxypropanol as the alcohol metabolite.	The chemicals within the category are defined as esters of phosphoric acid, having a phosphoric acid backbone with various alkyl substituents as illustrated.	http://www.epa.gov/hpv/pubs/summaries/phsacdde/c13356tc.htm EPA (2004c)
Cinnamyl Derivatives	The aromatic cinnamaldehyde derivatives are readily oxidized to cinnamic acid derivatives. The urinary metabolites of cinnamyl alcohol and cinnamaldehyde are mainly derived from metabolism of cinnamic acid.	The four substances in this group are un-substituted or alkyl-substituted cinnamaldehyde or 2,3-dihydrocinnamaldehyde derivatives. Common structural features among members of this chemical category are that these contain either a 3-phenyl-2-propenal or 3phenylpropanal backbone.	EPA (2005)
Benzyl Derivatives	The benzaldehyde derivatives are readily oxidized to the corresponding benzoic acid derivatives while the benzyl esters are hydrolyzed to yield benzyl alcohol that is subsequently oxidized to benzoic acid as a stable metabolite or endproduct. The benzoate and 2-hydroxybenzoates esters are hydrolyzed to yield benzoic acid and 2-hydroxybenzoic acid derivatives, respectively. The benzaldehyde derivatives are readily oxidized to the corresponding benzoic acid derivatives while the benzyl esters are hydrolyzed to yield benzyl alcohol that is subsequently oxidized to benzoic acid as a stable metabolite or endproduct. The benzoate and 2-hydroxybenzoates esters are hydrolyzed to yield benzoic acid and 2-hydroxybenzoic acid derivatives, respectively. As a stable animal metabolite, benzoic acid derivatives are efficiently excreted primarily in the urine.	The 10 substances are placed in the same category because all contain a benzene ring bonded directly to an oxygenated functional group (aldehyde or ester) that is hydrolyzed and/or oxidized to a benzoic acid derivative."	EPA (2002b).

3.2.3.4 Mode/ mechanisms of action or adverse outcome pathways (MOA/AOP)

In analysing the elements of read-across justification, mode/mechanistic understanding is a key element. With each grouping a description of the likely mode or mechanisms of action is specific and should be considered together with its limitation and purpose.

The ability to predict/fill a data gap of a target chemical is often affected by the mechanistic complexity of the toxicity endpoint. In general, endpoints with simpler mechanisms (e.g., sensitisation, mutagenicity) can be more easily predicted than those with multiple mechanisms. Also, values such as NOAELs are actually composites of various toxicity endpoints with the lowest figure arbitrarily selected. It will likely be difficult to interpret trends in these composite endpoints.

The mechanistic basis of developing a category including modes and/or mechanisms of action[40] is described in Chapter 2, specifically Section 2.4.3 that includes discussion of the development of adverse outcome pathways (AOPs).

If it can be demonstrated that the mode or mechanism of action for the toxicological or ecotoxicological effect is the same for similar structures or functional groups, then the confidence of the read-across from a source to a target chemical is significantly increased.

Within toxicology, there are a number of commonly held modes of actions for different endpoints, developed over a period time for different classes of substances. A proposed mode of action can take time to gain scientific consensus regarding its validity due to its complex nature (e.g., PPARα agonist-induced rodent tumours), while others are self-evident such as irritancy due to pH effects. To use a mode of action argument in support of a category, there needs to be consensus that it is a suitable and valid approach.

With the increasing availability of mode of action information, AOP and high throughput screening (HTS) as well as other predictive data, more integrative approaches can be explored to develop/ support hypotheses/justifications for read-across and selection of the most appropriate molecular descriptor(s).

Table 7. Examples of categories with similarities in mode/ mechanisms of action

Category	Mode/ mechanisms of action	Common structural features	Reference
Mononitroanilines	Toxicity is characterized by the ability to form methemoglobin in both humans and animals.	The chemicals selected for inclusion in this category are isomeric forms of the same base chemical	EPA (2003a)
Fuel Oils	The aquatic toxicity of products in the category are expected to fall within a narrow range regardless of the varying carbon number range and constituent composition of	The category was developed by grouping 8 ethylene industry streams made up of hydrocarbons that are generally carbon number 8 (i.e. C8) and higher with varying amounts of	EPA (2003b)

[40] defined in chapter 2

Category	Mode/ mechanisms of action	Common structural features	Reference
	those products, because the constituent chemicals of those products are neutral organic hydrocarbons whose toxic mode of action is non-polar narcosis. The mechanism of short-term toxicity [fish] for these chemicals is disruption of biological membrane function, and the differences between toxicities (i.e., LC/LL50, EC/EL50) can be explained by the differences between the target tissue-partitioning behaviour of the individual chemicals.	lower boiling materials. The streams are similar in that these are all complex streams that consist predominantly of the same higher-boiling hydrocarbons	

3.2.3.5 Chemical / biological interaction

The interaction of a chemical at a molecular target leading to a particular adverse outcome, also called the molecular initiating event (MIE), can occur via different mechanisms. Many MIEs are defined in the form of covalent binding to proteins and/or DNA. These types of MIEs are based on the principles of organic chemistry (i.e. electrophile-nucleophile reactivity). In contrast, 'receptor binding' or binding to enzymes are often based on non-covalent interaction, which are more selective in nature. Within the AOP concept the MIE represents a primary anchor as the beginning of the cascade which can be linked to the intermediate key events leading to the specified final adverse effect.

If it can be demonstrated that the chemical / biological interaction of two or more substances with the same functional group(s) is similar, then the data from a source chemical can be used to read-across to the target substance for specific and defined endpoints. As such information on how the chemical interacts with biological (macro) molecules will allow for an initial description of the molecular structure limitations for chemical category members acting in a similar manner, the potential effects of toxicokinetics should be considered.

For several endpoints, such as skin sensitisation or mutagenicity, knowledge about chemical reactivity provides important information when applying the analogue and category approach. For skin sensitisation, one of the necessary steps 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 gene mutagenicity but where DNA represents the nucleophile. Structural alerts as encoded in Toxtree, TIMES and the OECD QSAR Toolbox could be used to characterize the specific reaction mechanism leading to covalent binding to protein or DNA. Experimental "in *in chemico*" systems could also be used to quantify the electrophile-nucleophile reactivity and to confirm the predicted reaction mechanism to support a read-across for e.g. skin sensitization (Aptula *et al*., 2006) or mutagenicity (Benigni *et al*., 2005).

Supporting data that could be examined to determine similarities or trends in reactivity could also include gene expression arrays, metabolomics and/or data from specific tests designed to support mode of action hypotheses for category members.

Mode and mechanism of action concepts can also facilitate the read-across for aquatic toxicity. The term "mode of action" is understood in a broader sense than "mechanism of action", the first being seen as an integrator of the general type of interaction of a chemical with the organism, while the second is perceived as the precise (bio)chemical molecular interaction related to the Molecular Initiating (MIE) or Key Event (KE) of an Adverse Outcome Pathway. According to one of the earliest classification schemes (Verhaar *et al*., 1992), four modes of actions are distinguished for acute aquatic toxicity: inert substances, relatively inert substances, reactive substances, and specifically acting substances. The toxicity of the substances in the first two groups (later known also as non-polar and polar narcotics, respectively) is mainly hydrophobicity driven, while the second two groups (i.e. the reactive chemical substances and the specifically acting substances) form specific domains, and read-across between such domains is not trivial. The more precise definition of the mechanisms of aquatic toxicity can further facilitate the filling of data gaps. Some authors distinguish, instead of between reactive and specifically acting substances in relation to fish, between uncoupling of the oxidative phosphorylation, respiratory inhibition and electrophilic/nucleophilic mechanisms, electrophiles/proelectrophiles, acetylcholinesterase inhibitors, or central nervous system seizure agents (Russom *et al*., 1977). Other authors split further the electrophilic reactivity in specific reactivity mechanisms such as Michael type-addition, Schiff-base formation, etc. (Schultz *et al*., 2005). Different types of models could be used within a specific mechanistic domain (Netzeva *et al*., 2008). For substances within the same reactive mechanism of action, the potency of protein binding as predictor for e.g. acute aquatic toxicity, can be estimated in (semi-)quantitative manner (OECD QSAR Toolbox ver. 3.2).

3.2.3.6 Responses found in in vitro methods

There is increasing use of methods using molecular screening such as *in vitro* assays, gene arrays and "omics" analyses (see Chapter 2, Section 4) to inform biological endpoints/activities of chemicals. Many new assays are being used to screen large numbers of chemicals for particular effects, or used to characterise the intermediate key events within an AOP.

The similarity of responses measured in these *in vitro* assays between potential source and target chemicals for specific endpoint(s) may enable confidence in the read-across for related endpoints where there is further data from the source chemical in a more standard/traditional assay (e.g., *in vivo* toxicity data).

In order to provide solid mechanistic reasoning to use in vitro methods, it is useful to have transparent descriptions of a plausible progression of effects at the different levels of biological organization provided by AOPs. As explained in section 2.4.3.1, the AOP approach is a bottom up approach where events measured at the *in chemico* and *in vitro* level are linked to events measured at the *in vivo* level. For example in fish, estrogen agonists bind to the estrogen receptor, which can be measured *in chemico*[41] and *in vitro*, and set off a cascade of responses including the up regulation of vitellogenin production in the liver, which can also be measured *in vitro*, the conversion of testes to ova and the feminization of males observed *in vivo* leading to reproductive impairment and a decrease in the population.

The incorporation of *in vitro* methods into any integrated testing or assessment scheme allows for the employment of relatively rapid and often, but related to the complexity of the AOP and number of *in vitro* tests needed to cover all significant MIE(s) and KEs, inexpensive

[41] *in chemico* - refers to an abiotic measurement of chemical reactivity

hypothesis-driven testing. The same goes for employment of *in chemico* methods – and for predictive models building upon training set data generated by use of such in vitro *and in chemico* methods In such a scenario, the hypothesis that the target and source chemical have similar adverse effect on an apical endpoint can be tested by applying appropriate *in silico*, *in chemico*, and *in vitro* methods identified from the integrated scheme and depending on their reliability.

3.3. Trend analysis and computational methods based on internal models

For a given category endpoint, the category members are related by a trend such that the properties of the category members change in a predictable manner and there is a pattern in the changing potency of the properties across the category. The trend could be related to molecular mass, carbon chain length, or to some other physical-chemical properties or other molecular descriptors or profilers which are plausibly related to the AOP. For example a category with increasing chain length, with a common functional group, will affect solubility / log K_{ow}, which in turn may affect bioavailability and hence toxicity, both mammalian and aquatic. Analysis of these changes is referred to as trend analysis.

For larger categories, it is possible that several different relationships can be established for a single endpoint, defining subcategories. However, when developing such an approach, not all properties may change in a linear manner with incremental change in a structural element within the category.

A chemical that identifies a turning point in a trend is called a breakpoint chemical. Category members falling at the opposite extremes of a trend and between which interpolations are considered reliable are called sentinel or boundary chemicals.

Computational methods exist to perform a trend analysis and define a local (Q)SAR for a defined category. In particular the OECD (Q)SAR Toolbox allows application of a trend analysis to a series of related chemicals that belong to a formal chemical category or to a manually defined category composed of a few selected analogues. Separate guidance on this topic is available[42].

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 may be 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. However, it should be noted that a trend, when increasing or decreasing, is an expression of regression function and sensible statistical parameters should be demonstrated to justify that the trend actually can be used for predictive purposes.

The observation of a trend in the experimental data for a given endpoint across chemicals can be used as the basis for interpolation and may also be acceptable in certain cases for extrapolation (see Figure 1). Interpolation is the estimation of a value for a member using

[42] OECD (2009) Guidance Document for Using the OECD (Q)SAR Application Toolbox to Develop Chemical Categories According to the OECD Guidance on Grouping Chemicals, No. 102, Series on Testing and Assessment, ENV/JM/MONO(2009)5, OECD, Paris. (Note - The 2009 guidance makes reference to version 1.1 of the OECD QSAR Toolbox. A more recent version of the QSAR Toolbox is now available, but the principles are identical.)

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 can be performed when the series of values is monotonic (all increasing or decreasing) or when data are non-monotonic (e.g., parabolic). However, even in such circumstances, a substance that is not covered by other members can break the trend and show different effect. Sometimes if the level of confidence in the prediction is too low, the prediction may not be attempted.

Interpolation between category members is preferred to extrapolation, because it is plausibly more certain to employ. However, it may, in certain cases, be possible that data are available for a significant number of members of a category but are not available for a boundary chemical. In this case extrapolation to the boundary substance(s) may be considered as in an analogue approach, with its own justification. The potential for greater uncertainty in applying the analogue approach, then, should also be addressed.

Although the category approach is most robust when a quantitative trend between the category members can be established, it is theoretically possible to predict the presence or absence of a property or effect by applying trend analysis. Nonetheless, a lack of observed toxic effects for a chemical substance in a study for 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, careful evaluation of data. 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 the 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, but rather bioavailability of certain chemicals in the category may be affected (e.g., maximum bioaccumulation at some value of hydrophobicity and lack of other mechanisms for accumulation than passive diffusion). 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.

The important aspect to demonstrate is whether properties change as hypothesised in a predictable fashion with the incremental changes in the category. For example, it is important to provide evidence that the absorption is actually lower as the molecular weight increases, or that decreasing water solubility and logKow affect bioavailability and hence potency of aquatic toxicity in a series.

A trend might also be expressed as a quantitative activity-activity relationship (QAAR). 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) and only the sensitivity differs. It should be noted, however, that this concept seems better applicable to aquatic toxicity endpoints than for health endpoints.

Thus, a chemical category can be seen as a set of "internal" (Q)SARs (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" (Q)SARs (or QAARs) because these are derived directly from the experimental data for the category members. These models are also likely to be "local" models in the sense that these are based on a defined data set. Such an internal local model was developed for acute aquatic toxicity for the category of long-chain alcohols (C6- primary aliphatic alcohols) assessed within the OECD HPV Chemicals Programme OECD Cooperative Chemicals Assessment Programme(http://oecd.org/env/ehs/risk-assessment/oecdcooperative chemicalsassessmentprogramme.htm).

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 physical-chemical 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., Veith, G.D. (2007)) because it is a surrogate for the thermodynamic activity of the chemical in the blood and tissues; but acute inhalation toxicity is not directly related to carbon number or molecular weight because the degree of branching may be significantly different among the category members. Therefore, an approach using carbon number would not produce defensible extrapolations within this category whereas vapour pressure is a more reliable parameter to extrapolate the results from measured values to missing values.

3.3.1 Examples of trend analysis and breakpoints

To some extent all categories that have been proposed within OECD and regulatory fora have to display some degree of trend, whether it is increasing, decreasing or non-changing, in order to provide the justification for the grouping of those substances and any subsequent read-across. If a trend cannot be demonstrated in a category, the coherence of the category can be questioned. It should be noted that in some cases trends will be difficult to establish (e.g., in one-to-one read-across). In such case, the read-across should be justified by structural similarity and strong mechanistic considerations. Supporting evidence should be collected to strengthen the justification.

<u>Experimental basis</u>

A break point was noted in the aquatic toxicity of Long chain alcohols [C_6-C_{22}] as documented in the OECD high production volume chemical category (OECD, 2006b). Other breakpoints have been documented, for example the sensitisation potency of cinnamic

aldehydes (Patlewicz *et al*., 2001 as cited in ECETOC, 2012), was impacted by the log K_{ow}. Longer chain cationic surfactants were found to exhibit reduced eye irritancy (Patlewicz and El-Deredy, 1999 as cited in ECETOC 2012).

The above examples highlight the importance of documenting the trend exhibited by the category members. These also show how extrapolation through a category from, for example, low to high molecular weight may not always be appropriate unless other supporting data is available to justify the break in trend

Computational basis

Examples on how the OECD (Q)SAR Toolbox can be used for filling a data gap using trend analysis and for defining an internal model can be found on the OECD Quantitative Structure-Activity Relationships Project [(Q)SARs] pages[43].

The data for a particular endpoint can be used to construct a (Q)SAR that describes the properties of the members of the category.

An example of a (Q)SAR is the prediction of acute toxicity to an invertebrate species (*Tetrahymena pyriformis*) by means of a regression equation with the partitioning behaviour (log K_{ow} value) of the chemical as a descriptor (Schultz *et al*., 2002).

3.4. Computational methods based on external models

"External model" is used in distinction to the "internal model" described above and can refer to any model ((Q)SAR, 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, these could be based on experimental data that are obtained from a wider range of chemicals than those used in the category. Such external models can be "local models" for a congeneric series of compounds which is broader than the considered category or these can be as "global models", i.e. models based on a large and diverse set of training chemicals. The validity of the "external" (Q)SARs should be assessed according to 5 OECD (Q)SAR validation principles and it should be assessed whether the target substance lies within the applicability domain of the model. Some expert systems apply a combined approach in which the substances might be first split using some chemical or mechanistic rationale, and models are developed for the subgroups. It should be noted, however, that it cannot be expected that (Q)SARs are available, or can predict all types of substances and all types of endpoints. For complex health endpoints the read-across technique might be more informative than employment of statistical models if this is not supplemented with a detailed explanation about the predictions made.

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 predicted value should be compared with the experimental value available for other members of the category or an appropriate analogue. For example, a parabolic (Q)SAR could be used to characterise the trend in bioconcentration factor (BCF) values across a series of substances of increasing molecular weight.

[43] OECD Quantitative Structure-Activity Relationships Project: http://www.oecd.org/env/ehs/oecdquantitativestructure-activityrelationshipsprojectqsars.htm.

In other cases, model predictions may be used to identify additional analogues and rationalise, per endpoint, the category members that deviate from a trend. For example, a (Q)SAR or expert system might indicate that certain substances in a series have anomalous behaviour due to metabolism. Such an analysis should 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. SAR expert systems may also be a source of structural analogues for selected endpoints.

3.5. Reporting on incremental changes, trend analysis and computational methods

The nature of the incremental change should be documented as should any hypothesis which uses the information within the category. It should be explained which data within the category supports the hypothesis, especially if it is to be used as a means of data gap filling. The possibility of any breakpoints should be addressed.

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.

When making a prediction using a model, there are formats available providing information to facilitate regulatory consideration of both the model used and the prediction made. These formats were developed by the European Commission and are publicly available[44]. The (Q) SAR Model reporting Format (QMRF) follows the OECD principles for the validation of (Q)SARs[45]. The (Q)SAR Prediction Reporting Format (QPRF) enables the presentation of information necessary to assess robustness of the individual prediction.

For each endpoint addressed for read-across, a slightly different set of considerations will be required in order to justify the use of the read-across. Table 14 in the Appendix is an attempt to capture some of these specific points that should be considered when developing the read-across justification. Table 14 also lists available tools that may be of use in that process, and Table 15 provides references for these tools.

The process of developing category hypotheses for an analogue or category read-across and subsequently assessing its adequacy is given in Chapters 4 and 5, respectively.

However in general, the elements discussed in Section 3.2.3 are collated to form an overall hypothesis on the analogue or category approach. Multiple elements increase confidence in the category especially when it is driven by more detailed evidence such as when a common mode (and/or mechanisms) of action have been identified.

This hypothesis can then be tested with the available data on an evidence-based basis. In any such assessment, all the available data needs to be documented and assessed as to its

[44] EC DG JOINT RESEARCH CENTRE (2008) Institute for Health and Consumer Protection QSAR Prediction Reporting Format (QPRF) (version 1.2, September 2008) http://ihcp.jrc.ec.europa.eu/our_labs/computational_toxicology/qsar_tools/qrf/QPRF_version_1.1.doc

[45] OECD Quantitative Structure-Activity Relationships Project [(Q)SARs]: http://www.oecd.org/env/ehs/oecdquantitativestructure-activityrelationshipsprojectqsars.htm.

bearing on hypothesis and the potential read-across justification. This weight-of-evidence approach is inevitably case specific and details will vary considerably between different types of substances evaluated. Uncertainties should be highlighted, along with any actions that may be taken to reduce them.

There are a number of factors affecting the overall validity of the read-across which need to be addressed in any overall assessment. The first factor is unambiguity and clarity of chemical structure and composition (to the extent possible for UVCBs).

These include the quality and availability of the *in vivo* data used in the read-across for the endpoint of concern. With legacy information from different laboratories, test methods or test guidelines, and quality assurance / GLP status, the nature and quality of this data is likely to impact upon uncertainty and overall acceptance. Consequently it is necessary to consider data quality in the read-across justification.

The weight-of-evidence 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 other cases, such data might be used instead of experimental data. In an iterative process it may also be necessary to provide additional experimental data to gain sufficient confidence in the read-across justification so that extensive studies do not need to be conducted on the category. Once any new information becomes available it needs to be incorporated into the overall WoE conclusion.

External (Q)SAR predictions, if valid, could generally be included in a weight-of-evidence approach even if experimental data are available, especially when experimental data are of limited reliability or conflicts with each other, or for difficult-to-test substances. However, this does not mean that a read-across between (Q)SAR predictions should be used for data gap filling (in case there is no experimental data). Adding (Q)SAR predictions to experimental results is particularly useful if it may help in suggesting a mode of action of the chemicals assessed.

Guidance on weight of evidence considerations is provided in the OECD *Manual for the Assessment of Chemicals* (OECD, 2011b).

The use of data from studies that do not demonstrate any toxicological or environmental effect (often referred to as negative data) often need further corroboration compared to data from studies that show clear, significant effects. Although this inevitably requires additional effort for the data evaluation, such data are in fact valid for the read-across with corroboration. Corroborating data to ensure that the read-across from studies that showed no toxicological or environmental effect are in fact valid for the read-across. Such data may come from one of the other read-across elements detailed in the preceding sections. For example, a demonstration that the absorption and metabolism are similar between the source and target chemicals and that these follow a similar mode of action would be one way to demonstrate that the read-across in such cases may be appropriate.

4 ANALOGUE APPROACH: A STEPWISE PROCEDURE FOR IDENTIFYING ANALOGUES AND READ-ACROSS

4.1 Introduction

This chapter provides guidance on how to fill data gaps as appropriate for 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-standardized approaches. There has been considerable experience gained in the application of read-across using the analogue approach in the classification and labelling group of the EU. (ECB, 2005, Comber and Simpson, 2007, Gallegos Saliner *et al*., 2007, Hart, 2007, Hart and Veith, 2007, Schoeters and Verougstraete, 2007). Additional experience has been gained in risk assessments of existing chemicals (ESR programme; Tsakovska and Worth, 2007), and in the notification of new substances (NONS programme; Hanway and Evans, 2000). Recently, based on a learning-by-doing exercise, ECHA published an illustrative example of read-across using an analogue approach where advice to REACH registrants is given on how to prepare a case that would likely be satisfactory for REACH and CLP, ECHA (2013a).

Within the OECD Cooperative Chemicals Assessment Programme, read-across has been extensively performed since 1998. Examples of initial hazard assessments that rely on data from analogues, and that have been published, include: isobutanol (CAS No 78-83-1), p-chlorotoluene (CAS No 106-43-4), pentanoic acid (CAS. 109-52-4), 2-ethylhexyl acetate (CAS. 103-09-3), n-propyl acetate (CAS. 109-60-4), N-cyclohexylbenzothiazole-2-sulphenmide (CAS. 95-33-0), and several organosilanes such as methyltriacetoxysilane (CAS No 4253-34-3), trimethoxymethylsilane (CAS. 1185-55-3). These initial assessments are available from the OECD Existing Chemicals Database.

(www.oecd.org/env/existingchemicals/data)

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 for which data are available (e.g., OECD Cooperative Chemicals Assessment Programme, OECD Member Countries, or the EU Existing Substances Programme, and more recently REACH) 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 and Worth, 2007). 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 (multi-constituent substances and UVCBs) where there are dominating constituents, read-across by non-standardized approaches often involves the identification of a chemical substructure that is common to the target substance and its analogue(s) (or their respective breakdown products) and one of the following assumptions:

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

- In the case of quantitative read-across, the known value of a property for the analogue (source substance) can be used to estimate the unknown value of the same property for the substance of interest (target substance). In the case of a toxicological effect (human health or ecotoxicological), this assumption implies that the potency of an effect shared by the two substances is similar; or
- In the case of complex substances, the basis for comparison is likely to be different. For example, complex substances derived from certain process streams having similar composition may, to a large extent, 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, e.g., biological properties and bio-activation processes. A recent publication by Wu *et al.,* (2010) described a framework for identifying analogues and evaluating their suitability for filling data gaps. Analogues are categorized to reflect assumptions and uncertainties inherent in their use. Metabolism evaluation, similarity between source and target substances, and knowledge of key biochemical processes leading to an effect (i.e., adverse outcome pathway) play an increasingly important role in the identification of suitable analogues and in making predictions for the target substance. This is discussed in further detail in Chapter 3.

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. Figure 2 provides an illustration of this approach.

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

Before considering whether to use an analogue approach, the first step should be to determine whether the chemical is a named member of an existing category. Information sources on existing categories include:

- US EPA: http://cfpub.epa.gov/hpv-s/
- Canada: http://www.chemicalsubstanceschimiques.gc.ca/plan/index-eng.php
- OECD: www.oecd.org/env/existingchemicals/data
- eChemPortal: www.echemportal.org
- OECD QSAR Toolbox: http://www.oecd.org/env/ehs/risk-assessment/theoecdqsartoolbox.htm.

4.2.2 Step 1: Identification of potential analogues

There are a number of different ways to identify potential analogues as source substances with data with which the target substance can be compared. In cases where there is no presumption/restriction of what analogues to use, one can rely on a number of tools and techniques to assist and facilitate the identification of analogues. Tools and techniques for identifying analogues have been discussed in more detail in two ECETOC Technical Reports (ECETOC, 2010; ECETOC, 2012) and are summarised briefly here. Some of these tools facilitate the identification of analogues with and without data, whereas others can be searched to find associated data on a substance by substance basis.

The identification strategy is an exploratory process, and is not intended to be an element of the read-across rationale. A systematic search strategy may identify additional potential analogues for comparison, and if a significant number of analogues are identified, then a wider category approach may be justified, as discussed in the next chapter.

In terms of the systematic search strategy, common analogue identification approaches still rely on structural similarity or sub-structural assessment. It is well established now that structural similarity is only one criterion used to identify and evaluate the suitability of analogues for read-across. Nevertheless, structural similarity can be a pragmatic step in identifying promising analogues that could be expected to exhibit similarity in activity. Common mechanistic views might also be a starting point for analogue research.

The most commonly used structural similarity approach takes the form of a similarity index, a quantitative measure between 0 and 1 that summarises the commonality in structure based on the presence and absence of particular structural fragments. By far the most common that is seen is the Tanimoto index, which is defined as follows.

$T = NAB / (NA + NB - NAB)$
Where:
NA is number of features (ON bits) in A;
NB is the number of features (ON bits) in B; and
NAB is the number of features (ON bits) common to both A and B.

Essentially, a Tanimoto similarity index of 1 indicates the same structure, whereas an index close to 0 indicates a complete dissimilarity.

The pharmaceutical industry, which is the predominant user of the concept of molecular similarity, employs such 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 physical-chemical properties (e.g., solubility, portioning).

Tools such as ChemIDplus, freely available as a service on the National Library of Medicine website (http://chem.sis.nlm.nih.gov/chemidplus), provide the means to search on the basis of Tanimoto similarity or substructural inventories of chemicals, many of which contain links to available databases or literature information. Other tools such as Leadscope (http://www.leadscope.com) , a commercial tool that enables a sub-structural or similarity search which can be filtered to present results for only those analogues that actually possess associated information that might be useful for read-across purposes, are also available. AIM, the US EPA's Analog Identification Methodology (http://www.epa.gov/oppt/sf/tools/aim.htm) , works on a different basis. Rather than using a scoring scheme such as a Tanimoto index, the set of fragments and structural features that are encoded in the programs that EPA already uses as part of its estimation toolbox are used as a means of identifying similar analogues but with associated data. There are other many searching tools, and a non-exhaustive list is provided in Table 6 at the end of this chapter for illustrative purposes.

One of the most extensively used tools now, is the OECD QSAR Toolbox (OECD, 2013). Conveniently packaged with inventories of chemicals and a number of different databases, it provides a means of identifying analogues with available data from many sources. There are different ways to identify analogues, from structural similarity approaches based on fingerprints to identifying starting sets of analogues with commonality in one of the many profilers within the Toolbox. For example, a substance may be profiled on the basis of functional groups or chemical classes. It may be profiled by rule bases, such as alerting groups for protein and DNA binding, or mutagenicity, or other toxicological endpoints. This enables a

search to be performed to retrieve analogues that might be more general in nature (e.g., structural similarity) or analogues that might be more specific to an endpoint of concern (e.g., mutagenicity).

The OECD QSAR Toolbox can be used to identify analogues using the techniques described above. The main features of the OECD QSAR Toolbox are:

- The identification of relevant structural characteristics and potential mechanism or mode of action of a target chemical;
- Identification of other chemicals (analogues) that have the same structural characteristics and/or mechanism of action; and
- Use of existing experimental data to fill the data gap(s).

The identification of structural characteristics and potential mechanisms or mode of action is achieved with a set of profilers. These profilers identify structural alerts involved in specific reactions or binding mechanisms relevant for different regulatory endpoints. The Toolbox can then scan its databases to find other chemicals that have the same profile. Among the chemicals or analogues identified, some will have experimental data that the assessor may judge to be relevant for filling the data gap. The list of profilers is included in version 1.1 of the Toolbox and examples illustrating the analogue search workflow are provided in Diderich (2010). This list has been expanded considerably in the latest version (Version 3.2 released in December 2013).

Guidance is available for using the OECD QSAR Toolbox to use the analogue approach to fill data gaps and develop chemical categories (OECD, 2009a). Guidance is available for using the OECD QSAR Toolbox to inform an analogue approach for filling data gaps and developing chemical categories (OECD, 2009a). It should be noted that the OECD Guidance Document published in 2009 makes reference to version 1.1 of the OECD QSAR Toolbox and that more recent versions are now available; however the principles are identical. Additional guidance documents are available on strategies for grouping chemicals for data gap filling for acute aquatic toxicity endpoints, skin sensitization, genotoxicity and genotoxic carcinogenicity (OECD, 2012a, 2013a).

The extent to which differences in the purity or impurities are likely to influence the overall toxicity needs to be addressed and, where technically possible, excluded (see Section 3.2.3.1.5). (Q)SAR tools such as those described elsewhere in this report can be used to give an indication as to whether these impurities could have biological activity (such as genotoxicity or skin sensitising potential) that could influence how read-across can be used within the category.

4.2.3 Step 2: Data gathering for the analogues

For the source analogues chosen, published and unpublished data should be gathered on standard physical-chemical properties, environmental fate, ecotoxicological and toxicological effects. The standard information required depends on the regulatory programme but generally includes physical state, MW, logKow and other partition coefficients (e.g., Henry's Law coefficient, soil organic-carbon partition coefficient), aqueous solubility, particle size and structure, vapour pressure, melting point, and boiling point. Since these physical-chemical properties provide basic information on environmental distribution, fate and bioavailability, these can often provide supporting information for the read-across. The data gathering should

include all existing relevant data. 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 (www.oecd.org/env/existingchemicals/data). Information on chemicals can also be searched via eChemPortal (http://www.echemportal.org), which provides free public access to information on properties of chemicals (i.e., physical-chemical properties, toxicity, ecotoxicity and environmental fate and behaviour properties). In 2013 the eChemPortal had information from twenty-six data sources. Four of these can be queried using multiple criteria in the *Chemical Property Data Search* window. The list of data sources participating in eChemPortal is not closed and sources are added on a regular basis.

4.2.4 Step 3: Evaluation of available data for adequacy

Data available from relevant peer-reviewed sources such as the OECD Cooperative Chemicals Assessment Programme, or from hazard and risk assessment programmes in OECD member countries or in the European Union, can normally be used for read across. However, expert judgment is still needed to evaluate the relevance or adequacy of the analogue approach and data gap filling; this is further detailed under section 4.2.5 below.

The available experimental data should be evaluated for adequacy. The OECD Manual for the Assessment of Chemicals provides guidance on assessing the reliability of experimental data (OECD, 2011b). A scoring system, such as the Klimisch scheme (Klimisch *et al*., 1997), should be used by the assessor to document his judgement of the reliability and adequacy of the data: a study conducted in accordance with international guidelines is usually considered suitable. Poor quality analogue data for a potentially good analogue would only result in poor prediction. In addition, information needs to be provided in enough detail to allow for an adequate assessment, e.g. at least a robust study summary with enough information about significantly important experimental details relating to the observations and results obtained.

If read-across data have not been based on the most current test methods, particularly careful consideration of the quality and suitability of a method is important (Hanway and Evans, 2000). Usually, collection and evaluation of available data is performed in a stepwise manner.

4.2.5 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 Chapter 7). The matrix should include the chemical of interest (target chemical) and the analogue(s) (source chemical(s)). If multiple analogues are identified, these 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.

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

The next step after finding analogues and assessing the reliability of available data is the justification of the analogue approach. This should be done not only based on the structural similarity, but also on similar physical-chemical properties or other molecular descriptors or profilers relevant for the AOP in question between the candidate analogues and the target chemical. Criteria such as possible key functional group, biochemical processes and mechanism of action, or environmental fate are also important to consider. Wherever possible,

the relevance of the read-across of other endpoints should be evaluated in the light of the known or suspected mode or mechanism 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 for which the mode of action is likely to be similar or can cautiously be assumed to be related.

There are several ways to screen analogue(s) for their suitability and find arguments to justify the approach:

- Databases with *in vivo* data and other systems enabling the classification of substances according to structure, functional groups, possible mechanisms of action (e.g., OECD QSAR Toolbox);
- Expert systems able to retrieve and combine data in a search strategy, a more systematic approach to judge the adequacy of selected candidate analogues; and
- The use of (Q)SAR 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.

Factors shown in Chapter 3, Section 3.2.3 need to be addressed when evaluating the results of a read-across using an analogue approach. The supporting evidence discussed in subsequent sections in Chapter 3 should also be considered.

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 constituents.

If the read-across is considered to be suitable, the missing data for the target chemical (s) are 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, the following three options are possible:

It may be necessary to identify an alternative analogue – the best analogue may indeed not have the relevant experimental data, so it may be necessary to choose an analogue of lower quality in order to obtain data;

The use of a more extended category approach can be considered; or

It may also be necessary to obtain the information directly by testing.

4.2.7 Step 6: Document the analogue approach

If the read-across is considered to be suitable, the approach should be justified and documented according to an appropriate format in order to describe the approach being 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 (see Chapter 3). Ideally examples of unsuitable read-across should also be documented.

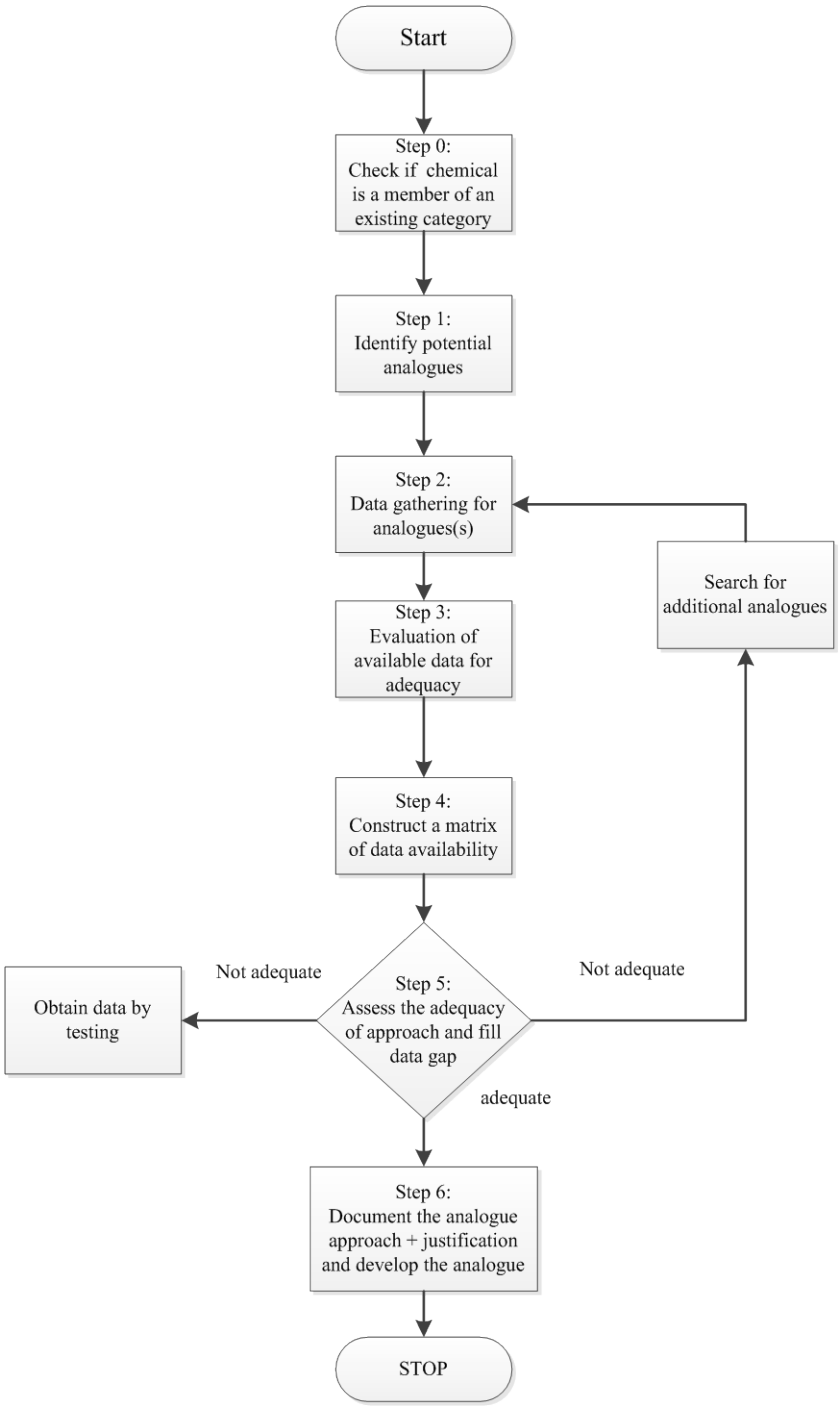
Table 6 - Selected example tools for analogue- searching

Tool & Website	Remarks
OECD QSAR Toolbox www.oecd.org/env/hazard/qsar	Freely available, the OECD QSAR Toolbox contains tools for systematic searching of analogues and

Tool & Website	Remarks
	databases of experimental results, as well as methods to form chemical categories and fill data gaps by read-across, trend analysis and (Q)SARs. Downloadable
AIM http://aim.epa.gov	US EPA's Analogue Identification Methodology. Links to publicly available, experimental toxicity data for target chemical as well as structural analogues. Web based version discontinued (2011) and a downloadable software version made available.
Ambit http://ambit.sourceforge.net/euras	Developed by IdeaConsult Ltd. Chemical databases and functional tools, including a tool for defining applicability domain of (Q)SAR models. Online use
ChemFinder http://www.chemfinder.com	Publicly available and subscription scientific databases. Online use
ChemID Plus http://chem.sis.nlm.nih.gov/chemidplus	Publicly available database from the US National Library of Medicine (NLM). Online use
ChemSpider http://cssp.chemspider.com/	Database containing more than 26 million unique molecules from over 400 data sources Publicly available for online use
Derek Nexus https://www.lhasalimited.org	A SAR-based system covering a wide range of toxicological endpoints in humans, other mammals and bacteria. Commercial
Distributed Structure-Searchable Toxicity (DSSTox) Database Network http://www.epa.gov/ncct/dsstox/	The DSSTox website provides a public forum for publishing downloadable, structure-searchable, standardized chemical structure files associated with chemical inventories or toxicity data sets of environmental relevance. Online use
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). Online use
Danish (Q)SAR Database http://qsar.food.dtu.dk	Publicly available version of the (Q)SAR prediction database developed by DK EPA. Online use – includes advanced search tools
Leadscope http://www.leadscope.com	Commercially available databases and (Q)SAR functionalities. Contains databases with experimental data.

Tool & Website	Remarks
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. Online use

Figure 2 - Stepwise approach to an analogue approach



5 CATEGORY APPROACH: A STEPWISE PROCEDURE FOR GROUPING CHEMICALS AND READ-ACROSS

5.1 Introduction

This chapter provides guidance on how to develop a category and fill data gaps as appropriate for one or more substances using the category approach. Chemical categories provide a useful framework for collecting available hazard information that is relevant to members of the category. If reliable hazard information is available, it can be used to assist in hazard classification and labelling decisions and/or for performing hazard and risk assessments for all category members that were justified, thus obviating the need to conduct extensive testing

The review of the use of chemical categories that was carried out in preparation for the development of this guidance[46] identified the following lessons learned with regard to the use of the chemical category concept:

- Initial hazard assessments that applied the chemical category approach were agreed upon by OECD member countries for 514 chemicals in 86 different categories as of 2012.
- Currently more than a third of the substances assessed yearly within the OECD Cooperative Chemicals Assessment Programme are assessed through the use of chemical categories, and this fraction is expected to increase significantly over the next few years as experience grows in member countries.
- As already concluded for the US HPV Challenge Programme, chemical categories can be used to estimate results for both environmental and human health endpoints. A recent retrospective analysis of the HPV voluntary programme concluded that participating chemical manufacturers filled 55% of health and environmental effects endpoints (that could otherwise have required animal testing) by applying read-across from animal tests already conducted or proposed for analogous chemicals (Bisschop *et al*., 2012).

The guidance in this Chapter documents a stepwise approach to the formation of categories. In the previous edition of the guidance on grouping (OECD, 2007), category formation was based on the use of non-computational methods. OECD guidance is now also available for using the OECD QSAR Toolbox to develop chemical categories according to the present guidance document (OECD, 2009a), Chapter 4. Guidance is available for using the OECD QSAR Toolbox to use the analogue approach to fill data gaps and develop chemical categories (OECD, 2009a) and provides examples for category formation using the QSAR Toolbox. It should be noted however that the Guidance Document published in 2009 makes reference to version 1.1 of the OECD QSAR Application Toolbox and that newer versions are now available for aquatic toxicity, skin sensitization and genotoxicity (OECD, 2013a). The principles are identical. It is emphasised that such computational tools can supplement, but do not replace, the need for expert judgment, which is required throughout the process. While 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

[46] Modified from ECB, (2005)

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, hazard assessment (characterization) and risk assessment. However, given the large number and diversity of substances that exist, and the extensive number of categories that may be formed, guidance on how to develop and evaluate substance categories cannot be rigid. Rather, this section describes how information on chemical properties, activities and, when available, metabolism and mechanisms of action should be gathered and combined with expert judgment to form robust and well rationalised categories, as well as guidance on how to document the justification for each category. The experience gained in application of this approach will lead to further improvements.

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 justified (see Chapter 3), and to do this, certain types of information should be documented (see Chapter 7). In order to collect this information in a systematic and transparent manner, it is recommended to follow a stepwise approach (Figure 3). 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 possible to start from a single chemical, or small group of chemicals, and to identify analogues to establish a larger category. Alternatively, larger inventories containing relevant experimental data can be trimmed down to find suitable analogues. Sometimes, 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

The first step in forming a chemical category is the same as the first step in determining whether or not to use an analogue or category approach: determine whether the chemical(s) is (are) a named member of an existing category. As stated in Chapter 4, information sources on existing categories include:

- US EPA: http://cfpub.epa.gov/hpv-s/
- OECD: www.oecd.org/env/existingchemicals/data
- Canada: http://www.chemicalsubstanceschimiques.gc.ca/plan/index-eng.php
- eChemportal: www.echemportal.org
- OECD QSAR Toolbox: www.qsartoolbox.org

A number of industry sectors have applied the principles of “grouping” for use in assessment of health and environmental hazard properties. Examples include petroleum substances (Concawe 2001; IPIECA 2010), dyes and pigments (ETAD, 2000), chlorinated paraffins (CPIA, undated), surfactants (CESIO, 2000 and 2003) hydrocarbon solvents (HSPA, 2002), acrylate resins (UV/EB Acrylate Resins, 2003), petroleum additives (ATC, 2000a and 2000b), bitumen (Eurobitume, 2002) (see ECB, 2005), and certain metals and inorganics.

Categorisation approaches have also been applied to flavours and fragrances (Salvito, 2007) and to other chemicals JECFA, the US HPV Challenge Program, SPORT, and the safety

assessment of fragrance ingredients under the Research Institute for Fragrance Materials (RIFM). For some substances (e.g., metals, phthalates), a category approach was used to assist in the categorization and prioritization of substances on Canada's Domestic Substances List (Environment Canada, 2003), and a category approach has been used in formulating a number of groups of substances that are currently undergoing risk assessment.

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

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. See Chapter 3, Section 3.2.3 for a discussion on the elements that can form the basis of the grouping and how these can be reported. 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), or chemicals having a common moiety of interest following degradation. 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, if the chemical similarity is present. 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 carbon chain length, basic hydrocarbon structure – aliphatic, cyclic 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 non-branched, cyclic or branched aliphatics (IHSC, 2004 and 2005). Some categories have also been defined in terms of a metabolic pathway, i.e., these 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 assessed are shown in Chapter 6.

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. In particular, all the endpoints that can be expected to be relevant for the category should be included. In general, the hypothesis when starting the grouping is that all substances and all *a priori* defined endpoints that are carefully justified to be linked to the categorization approach employed are covered by the category approach, i.e., the conclusions will be valid for all justified members of the category in the absence of endpoint data for some substances within the category. When some members of the category present specific features (e.g., branching) known to result in different properties, for example, these may be metabolised differently, these substances will deviate from the general trend of the category. In such cases, the category approach will be limited to only those endpoints for which the data robustly demonstrate that the trend is followed, while for other endpoints, individual conclusions or conclusions for sub-categories should be derived.

The category hypothesis should also address:

- 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;

- The mode of action/mechanistic rationale that provides a basis and understanding of the read-across within the category;

- The specific instances of read-across and trend analysis (interpolations and extrapolations), and any specific computational methods that have been used; and

- 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 substances not explicitly included in the current definition of a category. In most if not all cases, the inclusion/exclusion rules are stringent and the category is limited to the substances that are part of the initially formed category. Addition of new members to the category will require reconsidering the category justification and analogue research.

Individual members of the category are identified depending on the basis for the category. In many cases, this is done 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 the OECD 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 substances could also be 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. Scientific justification for such considerations should be always provided.

For practical or scientific reasons, a group of substances can be claimed as a category for one or several endpoints. These categories are usually called “targeted” categories.

The formation of a category has in many cases also been dependent 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.

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 applied to large inventories. For example, 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 and Veith, 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. While 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. The OECD QSAR Toolbox is also able to load lists of SMILES and provide an overview of data availability as a starting point.

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 exploratory in nature and likely to be dependent 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 (i.e., bottom-up or systematic approach). Examples of tools that might help include expert systems such as the OECD QSAR Toolbox, Derek Nexus (LHASA Ltd, UK) or other tools such as Leadscope (Leadscope Inc., USA) or AIM (US EPA). A variety of computer-based analogue-searching tools have been summarised in Tables 13 and 14 in the appendix to this guidance. 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 *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.

Numerous countries are implementing the category and analogue approach in their assessments of substances. The Domestic Substance List in Canada provides a good example of how chemical categorization was used to prioritize a chemical inventory. In 2006, the Domestic Substance List was subjected to a categorization process, resulting in the inclusion of about 4300 chemicals on the priority list. 500 of these were assessed between 2007 and 2012, whereas 1100 of the substances were considered to be low priority, and 3200 will be assessed between 2012 and 2020 Environment Canada (2010). To tackle this large group of chemicals, Canada used the OECD QSAR Toolbox, structural similarity and profiling tools to form groups of chemicals for further assessment. In the next phase of assessment, aromatic azo and benzidine-based substances, substituted diphenylamines, methylenediphenyl diisocyanates and diamines, organic flame retardants, phthalates, and metal-containing substances will be evaluated.

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). Confirmation from reference databases (e.g. CAS Registry) might be needed, and the ultimate description is provided from a set of comprehensive analytical results. Checking these inventories could indicate which regulatory jurisdiction might have additional information on the chemicals being evaluated.

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. Where purity profiles differ, it is important to describe how these differences may affect the toxicity of the substance(s).

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 physical-chemical 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, if it is important to the category justification, or indicates hazard (e.g., bioaccumulation, metabolism and cancer studies are relevant but not part of SIDS in the OECD Cooperative Chemicals Assessment Programme). In some cases, estimated data can be included with the premise that these data add value to the assessment and 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 are available elsewhere (see the IUCLID Manual (EC, 2007). Reporting formats are also described in Chapter 7.

5.2.4 Step 3: Evaluate available data for adequacy

Available data should be evaluated for its adequacy using guidance such as the OECD Guidance for Determining the Quality of Data for the SIDS Dossier (see section 3.1 of the OECD Manual for the Assessment of Chemicals (OECD, 2011b)).

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 assessed 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 irritants 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 a positive effect taking into account the data from the whole category.

- Where the data suggest possible breakpoints, the data should be evaluated to ensure that these points 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 show 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 (in some OECD member states) compared to other lower alkanes) (Hoffman, 2008; Trimmer, 2008). 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. Findings, which do not support the category, should be excluded reported and interpreted because such outlying information may inform about the robustness of the category.

1. Logical steps for grouping chemicals using an AOP

When grouping chemicals using an AOP, it needs to be justified and documented that the chemicals in the group formed fit in the chemical space defined in the AOP (i.e., similarity in physical-chemical properties, presence of functional group, sub-structural fragment).

Furthermore, any information on chemicals in the group formed, which are shown to trigger the molecular initiating event , or key event(s) along the AOP, will contribute to the justification of the category. This can be done in various ways:

- For example a (Q)SAR may be used to predict the molecular initiating event (e.g., protein binding, oestrogen receptor binding), an in vitro/ex vivo assay may be used to support a molecular initiating event or a key event in an AOP (e.g., vitellogenin induction in fish liver slice; in vitro alterations in sodium flux through voltage-gated sodium channels, leading to neurotoxicity), in vivo data may support a key event specified in the AOP (e.g., similar specific histopathological findings or triggering of MOA related response such as VtG in blood plasma / plasma of male fish, change in sex of fish or, organ weight changes in rat, specific protein expression, specific animal behaviour).

The accumulation of evidence reinforces grouping chemicals together, even in the absence of all information for all chemicals along the AOP. The similarity of adverse outcome demonstrated in experimental studies on chemicals grouped together is also a justification that chemicals follow the AOP.

There are numerous examples available on practical applications of AOPs for forming toxicologically meaningful categories (Schultz, 2010), which include, for example, receptor binding pathways for phenolic oestrogen mimics, weak acid respiratory uncouplers, skin sensitization, etc. See Chapter 2 for a more in-depth discussion of AOPs.

5.2.4.1.1 Using mechanistic information as tools for justifying grouping of chemicals in a category

As indicated above, it is not needed to have laid out the full AOP from the initial molecular initiating event and key events through to the final adverse outcome before being able to build a chemical category around a common mode (and/or mechanism) of action or key event. Mechanistic information, such as chemical profiling, generally associated with bio-profiling information coming from molecular screening (e.g., high throughput *in vitro* assays, proteomic information) can be used to justify the grouping of chemicals around a given adverse outcome, provided a link can be established between the endpoint in the molecular screening and the adverse outcome. Some useful working definitions are provided in the glossary of this document.

Table 7 illustrates conceptually how mechanistic information may be relevant in forming grouping of substances. The table shows the various situations/scenarios when using mechanistic information in the form of profilers to justify the grouping of chemicals for a given apical endpoint. The full AOP description represents scenario 1; other scenarios are variations of available data along the pathway.

Table 7. Qualitative use of Mechanistic Information in Forming Chemical Categories

Scenario	MIE	KE1*	KE2	KEn	AO	Applications/Usefulness
1	●	●	●	●	●	most chronic effects
2	●	○	○	○	●	acute and some local effects
3	●	●	○	○	●	many local effects
4	○	●	○	○	●	some use**
5	○	●	●	○	●	some use***
6	●	●	○	○	○	no use
7	●	●	●	○	○	no use
8	○	●	○	○	○	no use

*: KE is defined as key event
**: some use when the KE is e.g., positive receptor binding and the adverse outcome (AO) is reproductive toxicity
***: some use when the KE1 is e.g. receptor binding, the KE2 is vitellogenin increase or decrease in fish, and the AO is reproductive toxicity.

Depending on the amount and the distribution of information/data (e.g., for several chemicals for an assay versus a few data points for an assay, or for several events along the pathway versus on a single event) utility may vary for a regulator to justify the chemical grouping for a particular use. For every event listed in Table 7 (e.g., KE1), there is the possibility of having data from one or several (typically less than 10) protocols or methods assessing that event. Having data for the same chemical evaluated in different assays allows for an evaluation or reproducibility of the event that is of value in assessing confidence in a

particular result. Conversely, having many rather than a few chemicals tested in a particular assay for an event is of greater value in assessing the confidence in the assay and the result. It is also more valuable for the category justification to have results from assays representing several different events than from several assays representing a single event. However, it is likely that relying solely on a key event that is at a high level of biological organization (i.e., a more integrative key event) is likely to run the risk of mixing chemicals where different mechanism lead to the same apical outcomes.

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 ideally 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 Chapter 7).

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:

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

A preliminary 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 and is largely a matter of expert judgment. 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 (i.e., analogue read-across, trend analysis, and the use of external (Q)SARs). If the initial assessment indicates that the category rationale is supported and that the category is sufficiently robust, data gaps can be filled according to the guidance in Chapter 3 and the chemical category can be finalised and documented.

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 log K_{ow} and based on a narcotic mechanism of action, the data gaps may be filled by data from a valid (Q)SAR for the category. Alternatively, the category can be sub-divided into a number of subcategories defined by the breakpoints in the category, and members assessed within each subcategory.

If the initial category is not sufficiently robust or justified, the following options should be considered:

- 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 or a chemical should be removed from category (return to step 1);
- 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; or

- 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 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, the following factors should be taken into consideration:

- The choice of test will be influenced by the results of the preliminary evaluation of the category (as well as any regulatory requirement); and
- If there are no data for any of the members of a category for a particular endpoint, testing of a limited number of carefully selected category members may be considered appropriate; and 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. For example, for a substance in a category where skin irritation is predicted, a simple *in vitro* test might be adequate for hazard identification and follow-up classification and labelling and risk assessment.
- When making the test plan it should also be taken into account that generally interpolation is preferred to extrapolation. Hence generally the substances defining the borders of the group should be tested if testing has not already been performed

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 under the OECD HPV Programme.

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, as described in the category reporting format discussed in Chapter 7.

The data matrix summarizes the existing data and is an important indicator of how the proposed testing will adequately characterise the category. Each endpoint should have a row 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 (e.g., to check for occurrence of potential break point in the trend). 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.

When selecting a sample to test, it should be representative of the substance manufactured or imported, including the presence of any manufacturing impurities. It should also 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 key substances for testing that are of little or no commercial importance. While 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 evaluation of the category

If new test data become available, the category should be revised and further assessed 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 may then be done to fill the appropriate endpoints for each category member. If there are sufficient experimental data to support the conclusion 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:

- Perform limited and targeted testing;
- Revise the category hypothesis (and therefore the applicability of the category in terms of members and/or endpoints); or
- As a last resort, abandon the category hypothesis and perform standard test according to the applicable regulatory requirements.

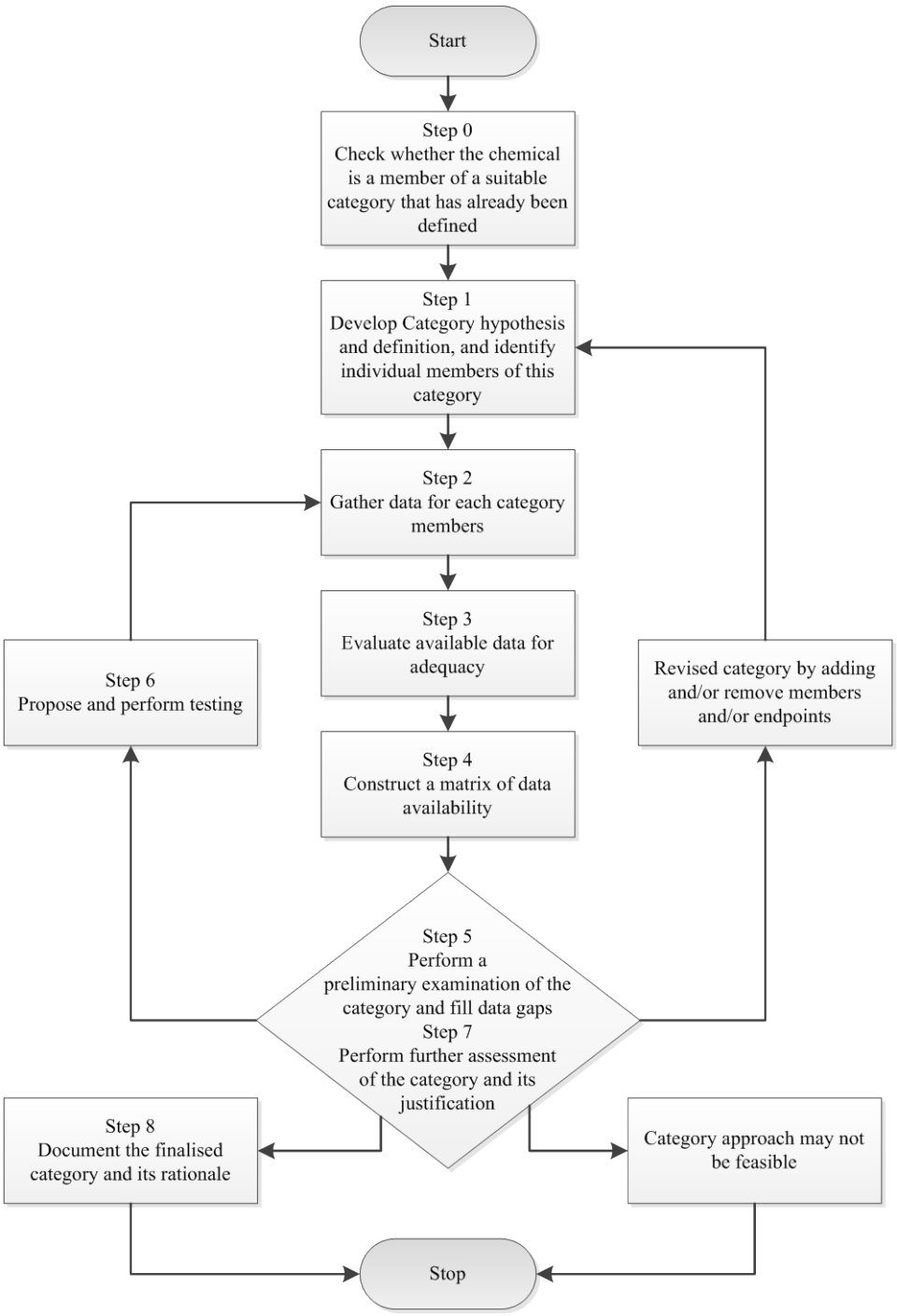
5.2.9 Step 8: Document the finalised category and refine the category rationale

The finalised category should be documented in the form of a suitable reporting format (see Chapter 7 for a proposed format). The category rationale should be initially elaborated and then refined when all available data have been examined. The category rationale should describe the common features of category members in terms of content (e.g. functional group, range of carbon number in the chain, and if relevant the degree of branching, cyclicity of certain members, the aromatic content) and if relevant in terms of fate (e.g., bioavailability, degradation and hydrolysis products, metabolism), and the reasoning for assessing the chemicals together as a group. In cases where the data do not support the category hypothesis, other information should be considered.

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.

Figure 3 - Stepwise approach to category development



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 examples and are not the only category types that might ever be formed or created. In addition, different regulatory frameworks might have different requirements for specific types of categories.

6.1 Chain length

Chain-length categories show an incremental, and usually constant, increase in chain length across the category. 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 Cooperative Chemicals Assessment Programme include alpha-olefins, higher olefins or monoethylene glycol ethers (UNEP Chemicals, 2006).

Categories defined by chain length generally show an incremental change in molecular weight and other physical-chemical properties, such as water solubility or log K_{ow}. However, not all properties will necessarily exhibit a linear relationship with chain length and care must be taken when making assumptions about such trends. For many homologous series, increasing log K_{ow} leads to increasing fish toxicity. While at the same time water solubility decreases. There is usually a point where the solubility is too low for toxicity to be expressed. For example, for the alpha-olefins category, 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 also be a change in 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. While 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, teratogenicity). 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 exemplified 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 subject to a particular regulatory programme but which could fit into the series should ideally 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.

(Q)SARs can be used to provide information for category members. In general, substances at either end of a chain length category should have all endpoints fulfilled, preferably with test data, if breakpoints have not been indicated. This would enable an interpolation of data for the other category members. For example, in the category on ethylene glycols, a linear regression

was used to predict acute aquatic toxicity, indicating that toxicity decreased with increasing chain length, and further supporting the low toxicity of the category members concluded from available experimental data (OECD, 2004b). 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 provide confidence in the inferences being made.

Other examples are oleochemical derivatives, which can be grouped in such categories as fatty acids or alkyl sulfates (OECD, 2007b). These categories may contain single-chain chemicals as well as complex substances containing chemicals of distinct chain lengths at varying amounts. The relative amounts of individual chain length molecules in complex substances are usually reflective of the chain length distribution in natural fats and oils from which these are derived. Since the category chemicals differ from each other only by the number of $-CH_2-CH_2-$ units, these categories are very often homogenous and exhibit a constant pattern in the changing of the potency of the properties across the category. However, great care should be given for the fact that the functional group introduced in the natural chemicals may change the metabolism or the mechanism of action.

6.2 Metabolic or degradation pathways and toxicokinetics

One of the rationales that can underpin a category is that of a common precursor and/or breakdown product that results via physical or biological processes (i.e., metabolic or degradation pathway similarity). The primary hypothesis that underpins a metabolic pathway similarity is that Substance A is metabolised to Substance B, C, D etc. Typically the metabolites most often considered are downstream blood metabolites though the discussion here may equally be applicable to other organ and tissue specific metabolites. The metabolic pathway approach may allow for the possibility of using data from the parent chemical to identify the hazards of the metabolites, or vice versa. Hazard identification studies with the parent chemical could be potentially 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. If the metabolism of the parent chemical within barrier tissue (e.g. lung, gut, placenta tissue) occurs so rapidly that the initial primary metabolite is the predominant chemical found within the blood, data from studies conducted with the primary metabolite itself can be used to characterise the hazards of the parent chemical. Examples of metabolic pathway categories that have been assessed within the OECD Cooperative Chemicals Assessment Programme include isobutyl isobutyrate (CAS No 97-85-8) or trimellitic anhydride (CAS No 552-30-7) (UNEP Chemicals, 2006). Other metabolic analogue approaches have been considered under REACH, an example is that of diethylene glycol phenyl ether presented at the ECHA-Cefic LRI read-across workshop see

(http://echa.europa.eu/documents/10162/5649897/ws_raa_20121003_metabolism_in_guidelines_echa_cefic_lri_workshop_ball_en.pdf), another example where the common metabolic pathway played an important role is on Glycol ethers E that was discussed as part of the ECETOC TR (ECETOC, 2012).

One of the first issues to consider when forming a metabolic series is to determine what metabolism information is available under the programme for which the chemical is being assessed. Toxicokinetics studies are not requested in many regulatory programmes and therefore would require a sponsor of the chemical to do additional work beyond what would be normally considered necessary. If metabolism information is not specifically required as part of the regulatory programme but is needed to substantiate the read-across hypothesis, *in vitro*

metabolic studies in conjunction with PBPK or PBPD models (the latter if feasible) may prove sufficient to characterise the metabolic pathway(s).

Ideally evidence would be collected from *in vivo* studies, where a parent chemical and primary and secondary metabolites in the blood would be measured directly. Gathering such information can be expensive, technically challenging and lengthy to run, particularly if radiolabelled test chemical is required, such that a toxicokinetic element in a standard toxicological study could be a potential alternative. ADME information could be conceivably generated as part of a range finding study. Saghir *et al.,* (2012) and Creton *et al.,* (2012) recently published overviews of the possible inclusion of toxicokinetic parameters in standard guideline studies from sub-acute to chronic repeated-dose toxicity studies, developmental and reproductive toxicity studies. The addition of toxicokinetic parameters to standard toxicological studies could be performed relatively inexpensively, does not require the use of radiolabelled material by default and measurements can be taken during the range finding study using a limited number of animals, thus not increasing the total number of animals used overall (ECETOC, 2012; Patlewicz *et al*., 2013). Nevertheless, it is noteworthy that the ADME study with radiolabeled substances provides more information.

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 that are not part of normal metabolism or require enzyme induction may be less well characterized. Even with metabolic pathways that are well characterised it may still be important to understand the kinetics of the metabolism. Thus a question remains over what are acceptable experimental data and justification to demonstrate the existence and rate of the metabolic pathway to support the use of read-across. There are no objective thresholds to specify what constitutes 'rapid' metabolism, 'predominant metabolite' or indeed whether the burden of proof should differ depending on the endpoint itself or the absence/presence of effects. The significance of toxicity of any residual parent chemical when reading across from metabolite to the parent may need to be considered. Exploring differences in toxicity and distribution (if any), where data on both parent and metabolite exist for the same endpoint can be helpful in this type of determination. A rigid framework of criteria would not be helpful, as each substance will be context dependent on its own available information. Practically, assessing the available toxicokinetic data (the generation of which has been discussed above) in the light of the entire toxicological dataset, including information on the mode of action of the chemical, should provide a measure of the validity in the metabolic pathway approach. Once a metabolic pathway has been demonstrated, if the toxicological profiles of the category members can be shown to be consistent, then the use of any read-across should be strengthened. If the approach was extended to include the consideration of other structurally related 'metabolic pairs' as part of a category approach then this would aid in reducing the uncertainty associated with the read-across proposed (Patlewicz *et al*., 2013).

The metabolic pathway approach is usually limited to systemic toxicity endpoints. 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 typically be addressed using the metabolic category logic. These sites of contact effects are often due to the physical-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 dependent 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. This type of information can be noted in the list of endpoints

covered by the approach as part of the documentation presented, i.e. the reporting format discussed in Chapter 7.

The following specific issues could be considered when developing a metabolic pathway category, according to the stepwise procedure described in Section 5.2.

Step 1: Provide definitive information on the metabolism of the parent chemical to the primary and secondary metabolite. This information may include time course data for either blood or tissue for both the parent chemical and the primary and secondary metabolites. In some cases *in vitro* metabolism data may be sufficient. Although in vivo data are recognised as more comprehensive (depending on study design), in vitro data and/or PBPK modelled data can be as useful and illustrative.

Step 2: The metabolism information should be examined to determine, if in fact, the primary and secondary metabolites are formed, if these achieve appreciable levels within the blood and/or tissues and determine basic toxicokinetic parameters for the parent material. If the metabolism of the parent chemical to the primary metabolite is rapid and is thought to occur within barrier tissues (e.g., lung, gut, placenta tissue), then it may be appropriate to use hazard identification studies from the primary metabolite to identify hazards associated with exposure to the parent chemical.

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 then structurally related metabolic pairs should be considered. If such information were not available, a study could be designed and conducted. In this case the parent compound could be tested. Any 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 levels of the parent chemical, primary and secondary metabolites.

Step 4: 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 can 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 steps above describe a situation where the parent chemical and its metabolites can be considered as one category because a parent is transformed into its respective metabolites. There is another possible scenario, in which a common transformation product is formed from several different parent chemicals. In this instance, toxicokinetic for all category members will ideally be needed to support the underlying hypothesis. Residual parent compounds, as well as possible by-products, may also need to be taken into account.

The metabolic pathway approach may not be applicable 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 directly from rodents to fish or between amphibia and fish or insects and other species due to the difference in the metabolic processes and enzymes present within those species. Also significant differences in metabolic capacity occurs often between different life stages and this should also be taken account of when relevant.

The same concept underlying the metabolic pathways can also 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, 2000b). A biotransformation/degradation pathway approach making use of biodegradation and metabolism studies can be useful to help in characterising bioaccumulation potential. If a parent substance was extensively degraded and/or metabolised, this type of information could be helpful to rationalise the likely bioaccumulation potential of a chemical. However it should be considered whether the biotransformation pathway and rate observed is likely to be relevant for the species in which the bioaccumulation is considered, because it is well known that biotransformation is often highly dependent on taxonomic group and life stage. It should for example be carefully considered how much standard biodegradation studies such as ready or simulation biodegradability studies including analysis of the transformation products formed can really provide pertinent information to substantiate the read-across proposed if this is considered relative to the BCF in fish.

6.3 Mixtures of discrete substances

Categories can be developed for series of chemical reaction products or multi-constituent substances (MCS) that are related in some regular fashion. In many cases, chemical reaction products are UVCBs. As with categories based on discrete substances, 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 Cooperative Chemicals Assessment Programme provide useful case studies on dealing with multi-constituent substances. Further information is available at (http://webnet.oecd.org/hpv/ui/ChemGroup.aspx) . 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 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 these were not commercially supplied (Caley *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 *et al*., 2007). Effectively, the ammonium ion was used as a marker for aquatic toxicity (see also Section 6.5).

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

6.4 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 substances in the mixtures are active in a common toxicity pathway, and so this approach is strictly only applicable for 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.

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 or MOA 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.

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 synergistic effects are conceivable, these 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 MOA- considerations.

In the Toxic Equivalents (TEQ) approach, the most toxicologically relevant 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).

$$\text{TEF (component A)} = \frac{\text{Reference effect value}}{\text{Component A effect Value}} \qquad \text{Equation1}$$

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):

Table 8. Toxic Equivalency Factors (TEFs) for dioxans and furans

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 the TEQ for the mixture is:

(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 risk assessment of coal tar pitch (under the EU Existing Substances Regulation, CAS 65996-93-2 Pitch, coal tar, high-temp, EC, 2008c 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.

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

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 doses of the individual components (discrete substances) taking their potency info into account, is additive. Furthermore, 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 mode of action 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.

6.5 Isomers

Isomers are chemicals that have identical chemical (or empirical) formula 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:

- Chain isomers. For example hydrocarbon chains with identical or variable lengths and variable branching patterns (see also section 6.1).
- Positional isomers. For example hydrocarbon chains with a functional group that varies in position along the chain. An example is 1-butene and 2-butene.
- Functional group isomers. These isomers also have identical molecular formulas, but contain different functional groups. Examples are butanal and butanone which both have the chemical (or empirical) 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. The following stereoisomerism can be distinguished:

- Diastereomers are non-superimposable stereoisomers: these are non-mirror images of each other. Cis-diastereomers have substituent groups projecting in the same direction; trans-diastereomers have substituents oriented in opposing directions. These diastereomers can occur when a double bond or a ring is present which restrict the rotation. 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: these 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 these rotate polarized light and how these interact with different optical isomers of other compounds, and

how these 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.

Tautomers are chemicals that have several possible arrangements of double bonds in a dynamic equilibrium. These arrangements may possess different biological and toxicological properties. For example enol-keto tautomerism in phenolic structures (e.g. hydroquinone) can change an apparent di-phenol to a quinone that is a more activated species of the same chemical structure. In order to predict reliably the behaviour of such chemicals, all tautomeric forms of the chemicals must be evaluated.

Stereoisomers can have similar or different chemical or toxicological properties. Even though these 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. Several illustrations of the impact of chirality on the toxicity and fate are given by Smith (2009).

The substance(s) with a data gap as well as substance(s) with data are similar such that their physical-chemical, biological, and toxicological properties would be expected to behave in a predictably similar manner or logically progress across a defined range. The incremental change is so small that it is not expected to affect the property sufficiently. This similar manner or logical progress should be demonstrated by the available experimental data. (Q)SAR models and trend analysis can also be used in addition of experimental data to support the estimate.

However, 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, these 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 inherently biodegraded. Therefore, it is not possible to assess the biodegradability of the inherently 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 (OECD, 2004c) and their mixtures, the following observations were made:

- 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.
- Extrapolating properties to isomeric mixtures should take into account mode of action, potential additivity and synergy, as well as purity profiles, and mixture composition.
- 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.

- 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.

6.6 Complex substances (UVCBs)

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, or UVCBs)". 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 be a cautious approach in defining categories and applying the following recommendations. There are many different types of complex substances, although generally these all have the following characteristics in common.

- These contain numerous chemicals and cannot be represented by a simple chemical structure or defined by a specific molecular formula.
- These are not intentional mixtures of chemicals.
- Many are of natural origin (e.g., crude oil, coal, plant extracts, reaction products) and cannot be completely separated into their constituent chemical species.
- The concept of "impurities" typically does not apply to complex substances (UVCBs).
- These are often produced according to a performance specification related to their physical-chemical properties.

While CAS numbers are important for identifying substances, in the case of complex substances these 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:

Petroleum substances are based on a hierarchy of considerations including chemical characteristics such as hydrocarbon type, carbon number range, content variability of aromatic, aliphatic naphthenic, aliphatics and S and N containing hetero-cyclic constituents. In addition, the source of the crude oil, production and processing characteristics such as distillation range as well as the last processing step have to be accounted for because this information provides essential insight in the characteristics of constituents potentially present.

- Coal derived complex substances are typically based on the applied production process and may include information on the distillation range and the chemical composition, and
- 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
- Complex inorganic substances are complex materials containing varying amounts of metals, metal compounds and/or minerals. These may occur naturally (e.g. mineral ores) or be manufactured during the various refining streams of the metal and mineral industry (e.g. refinables, metal intermediates). These often have no well-

defined CAS number, and their main identifiers are their name and their description. The description typically includes the origin of the substance, its production process (if applicable) and its main constituents.

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. Furthermore, some CAS numbers have a broad definition that may fit different, but related complex substances that fall into different categories. These complexities have sometimes led to the use of physical properties and chemical descriptors (e.g. chain length, chemical class, size of aromatic ring systems) as a way to define categories of complex substances. Recently, the OECD developed guidance on how oleochemical substances can be characterised in a way that their composition is accurately and consistently described for hazard assessment purposes (OECD, 2014a (in prep). In the case of NCS, this categorisation may also occur around the major chemical component(s) present, and might include marker chemicals for toxicity when it is clear that the behaviour of the UVCB substances are 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.

6.6.1 General guidance on developing categories for organic UVCBs

Complex substances, or substances of unknown or variable composition, complex reaction products and biological extracts or materials (UVCBs), pose challenges for hazard evaluation and for judging adequacy of read-across between similar complex substances for data gap filling. Therefore, the application of computational methods for generating representative structures, present or likely to be present in the UVCB, and to perform screening of their hazardous properties using non-test methods, is seen as beneficial. The application of these methods, however, does not intend to replace the experimental testing, as might be required according to different regulatory frameworks

The representative structures approach is mainly applicable to organic UVCBs, such as hydrocarbon solvents, as well as oligomers. Indeed, a tool (PETROTOX [47]) employing the Hydrocarbon Block Method (CONCAWE 1996) whereby a complex substance is divided into representative blocks of constituents with similar physical-chemical, fate and hazard properties has been developed. Inorganic and organometallic UVCB substances are however more difficult to handle in this way and other non-testing approaches or solutions are usually envisaged to address these.

The chemical representation and modelling of UVCB substances is inherently linked to the source, the manufacturing process and other identifiers, including analytical techniques and industry-specific identifiers and end-product quality indices, which can provide boundaries of the chemical space. The better the description of the UVCB substances especially in relation to chemical characteristics of its constituents, the more accurate a derivation of representative structures is possible.

The chemical space of a UVCB, which may encompass a large number of individual constituents, often prevents their full enumeration. The ability to generate representative

[47] CONCAWE https://www.concawe.eu/content/default.asp?PageID=778

structures and predict their hazardous properties can be used in hazard and risk assessment of UVCB substances by pointing out the regions of chemical space of greatest concern. Targeted analytical characterisation of constituents of representative structures can be determined in some cases, and the knowledge of constituent concentration(s) can inform (1) the category justification and read-across for UVCB substances and (2) the hazard identification of UVCB 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 repetitive changes in the constituents may include the length and branching of the hydrocarbon chain; presence and number of aromatic rings, presence, number and position of different functional groups, presence and position of heteroatoms, or different forms of isomerism.

There can be constraints arising from the manufacturing process, e.g. that the substance only contains distillates in a specified temperature range (cut-offs) or from performance indicators of the final product, such as a viscosity specification. Other physical-chemical identifiers, such as vapour pressure, flash point and self-ignition temperature may also be available and should be used, to the extent possible, for narrowing down the UVCB chemical boundaries. For inorganic UVCBs, similar constraints may be identifiable by geological, mineralogical and/or metallurgical experts.

The alkyl chain may differ in the length, the number of (conjugated) double and triple bonds, the degree of branching, the presence of aromatic and non-aromatic rings, and the position of the functional group(s). The position of unsaturated bonds may be limited to certain parts of the chain, as in alpha-olefins. Branching may also be limited to certain positions of the chain with respect to the unsaturated bond, such as the vinyl, allyl, or at carbon atoms further from the unsaturated bond. The alkyl branches may have odd, even or arbitrary number of carbons, depending on the source of the starting material, i.e. of natural or synthetic origin, and the process. The alkyl chain may be defined with a generic description, such as tallow, coco, or neo.

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

- Composition - for organic UVCBs, it is important to clearly characterise the identity of the constituents and the composition of the complex substance to the extent it is relevant for hazard characterization. A meaningful indication of variability should be provided. 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
- For biological NCS identification of the genus/species, origin should be considered
- If marker chemicals are appropriate, these should be clearly identified and if possible quantified for all category members

The critical issue when considering UVCB substances is composition. In order to determine the viability of using read-across one needs to understand the components of these products in sufficient detail. It is also necessary to determine which of these components are likely to drive potential effects (e.g. benzene, 1,3-butadiene, PAHs). It is recognised that generic criteria to describe the composition of UVCBs still need to be developed (OECD, 2012d). Criteria under development should consider concentration range and typical concentration of components of the UVCB; what are the generic constituents; what are the specific constituents; how to differentiate well-defined substances from UVCBs; what are acceptable constituent concentration ranges; and how to handle substances which are difficult to analyse in practice. For many UVCB substances, standard industry methods, including spectroscopic techniques (UV spectroscopy, IR spectroscopy, NMR, and mass spectrometry) and chromatographic techniques (gas, and liquid chromatography) that may provide adequate compositional information for some of the less complex UVCBs (e.g. hundreds of components) (Concawe, 2012), but for the more complex UVCBs (thousands, to hundreds of thousands, of components), there exist state-of-the-art techniques including two-dimensional gas chromatography (GCxGC), which can provide significant compositional insight into the substance. However, as is common with emerging technology, a high amount of effort and expertise is required to develop methodology, and also, importantly, to interpret the results (Concawe, 2012). As such, methodology does not currently exist for the analysis of complex UVCBs in a standardised manner.

Current recommendations are generic and include the use of standard industry methods for characterising UVCB substances of all sub-types, to allow for a structured analytical approach that allows for accurate hazard assessment. This discussion is still on-going at the time of writing (ECETOC, 2012).

- 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 which 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).

The systematic generation of representative structures through combinatorial algorithms is especially useful when screening for outlying behaviour of the constituents. Presence or absence of such constituents should be specifically checked and should be included in the description of the substance. The presence might also be included also in the name of the

substance, together with a quantitative indicator (e.g. > 2% aromatics). The screening in certain cases, like presence of alerting functional group, could be done on generic chemical description (e.g. by SMARTS[48] or Markush-type structure[49]). However, in most of the cases, full enumeration would be required. It is important to know the type of variations in structure that the complex substance could cover, and to avoid ignorance of potentially dangerous (classes of) constituents due to limitations in the multiplication algorithm. Some tools offer random combinations with high coverage of the theoretical chemical space, in case the number of the possible variants is too large. Therefore, tools that allow such enumeration need to be applied, followed by computational screening.

Toxicologically hazardous components might be present in negligible amount in substances. Similar overall composition does not necessarily mean toxicological similarity, since hazardous properties may arise from minor constituents. Therefore, the more detailed the identity information, the more precise computational analysis could be applied.

1. Data gap filling - Read-across/SAR and (Q)SAR for organic UVCBs

It is possible to fill data gaps within a defined category either using read-across/SAR or establishing a (Q)SAR, which is sometimes best described as a local (Q)SAR. 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 (Q)SAR 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.

In cases where no experimental data are available for one or more endpoint(s) of the category, or in cases where experimental data may be missing at the lower or upper boundary of the category, the use of data from surrogate substances not formally part of the category may be appropriate. Surrogate data may in particular be useful to reinforce a trend in the category and establish that there is no breakpoint within the category. This is the case when the surrogate data shows similar effects or a similar absence of effects as predicted in the category. An illustrative example is the use of a three –generation study on C9 aromatic hydrocarbon solvent to fill a data gap for the C10-C13 aromatic hydrocarbon solvents (www.oecd.org/env/existingchemicals/data).

In certain cases, there might be interactions between the constituents in the biological systems. Concentration addition is the default type of interactions. Nevertheless, independent action, and specific interaction (e.g. synergism/antagonism) could also appear. It is also very important to carefully consider the dose-response relationship for read-across/(Q)SAR versus the nature of the complex substances and the level of components of concern within the complex substances.

The computational demand in the multiplication and screening of organic UVCB components and the practical issues in handling these processes had led to a proposal for the

[48] SMARTS - Smiles arbitrary target specification, a line notation language for specifying molecular query patterns

[49] Markush structures (-R) are chemical symbols used to indicate a collection of chemicals with similar structures.

integration of the developed methodologies in a single software tool. Thus, the OECD QSAR Toolbox looks a promising platform for category building involving UVCB substances with known composition (and also multiconstituent substances, i.e. substances comprising more than one chemical structure) and for the development of a reporting format tailored to organic UVCBs.

2. *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.

3. *Good practices in developing categories for complex substances (UVCBs)*

In forming chemical categories made of complex substances, the following good practice should be observed to enable hazard assessment and regulatory acceptance of the proposed grouping and data gap filling:

- Distinguish the individual constituents as far as possible, providing compositions and variability ranges as narrow as possible;
- Name the substance in a clear and consistent way, which reflects both the constituents and the composition, in accordance with relevant and existing competent authority requirements;
- Define the substance in unambiguous way, with regard to inclusion or exclusion of (groups of) constituents;
- Support the identity and composition of the substances with analytical data to enable comparison with the category definition;
- Base the grouping on a scientifically credible and verifiable hypothesis;
- Analyse the constituents and substances for outlying behaviour;
- Explain, as for a normal chemical category how the read-across is being made for the various endpoints to fill data gaps;
- Sub-group the category, if there is a good reason for that;
- Identify reasonable worst case scenarios for chemical hazard by endpoint, for components in the substance, and for substances in (sub)category, and ensure these are reasonably addressed;
- Justify the mechanistic rationale for the category (per endpoint, if necessary);
- Before attempting hazard assessment, make sure that there is sufficient data to allow trend analysis and there is a good coverage of chemistry and properties within the (sub)category, otherwise consider testing;

The work on the grouping of complex substances in categories for read-across and data gap filling is still in its infancy. To approach grouping complex substances in a step-wise fashion, the following steps are proposed:

1. Provide all analytical, physical-chemical and manufacturing information on each complex or UVCB substance;
2. On the basis of the information in the above point, generate representative structures for all substances in the category, if possible;
3. Merge the representative structures for all substances in the category in one pool that covers the chemistry spanned by the whole set of UVCB substances in the category;
4. Collect all available information for the representative structures, including data from experimental databases or predictive methods, for the endpoint(s) for which there is a data gap;
5. Group the representative structures in groups that have similar hazard profile (or fate property depending on the endpoint that has the data gap);
6. Build an analytical matrix that shows the mass fraction of each UVCB substance for each identified group of representative structures; if concentration is unknown and there is significant variation in the profile of the different groups of representative structures, this should be evidence that the category cannot hold without further analytical characterisation;
7. Attempt to build a local (Q)SAR;
8. Model, using as independent variables, the analytical composition of the UVCB substances expressed as the mass fraction of the UVCB;
9. Or each group of representative structures that has a significant hazard profile, if possible.

Points 4-7 will need to be repeated for each endpoint. The category may stand for some endpoints but not for others (e.g. a category can stand for systemic but not topical effects, or *vice versa*. The computational approach depends also on the computational possibility to predict endpoints. Petroleum UVCBs are generally defined by manufacturing and processing conditions, hydrocarbon chemistry (e.g., aliphatic hydrocarbons, aromatic hydrocarbons), physical-chemical properties such as boiling range or carbon-number range, and common use categories. An example of the grouping of petroleum UVCBs, developed for the purposes of the former EU Existing Substances Regulation and also used for classification and labelling purposes, is given in Comber & Simpson (2007). According to this approach, petroleum UVCBs are grouped according to the process by which these are manufactured, on the assumption that substances within each group (or sub-group) have similar physical-chemical 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 (Q)SAR.

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

6.6.2 Hydrocarbon solvents

Hydrocarbon solvent categories are based on typical chemistry and carbon-number range. There is general agreement that the following identifiers should be provided in the composition of the category:

- CAS number,
- Carbon number range and percentage of the indicated carbon number range to be included under a given CAS number (e.g. C10-C13 >80%),
- Benzene content,
- Aromatics (including benzene),
- Boiling point range,
- Aliphatic content with regard to n-hexane content,
- Sulphur content,
- PINA distribution (Paraffins, Iso-paraffins, Naphthalenes and Aromatics).

There may be other streams with the same CAS number that may not be covered in the category assessment if these don't meet the identification criteria described in the assessment. 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.

6.6.3 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 physical-chemical properties.

6.6.4 Natural complex substances (NCS)

NCS can originate from plant, animal or microorganisms. Some inorganic UVCBs are also natural substances - e.g., natural clay minerals. For example, NCS include 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 (2007).

Ores and ore concentrates are naturally occurring substances. These are more complex and less well-defined than most substances, and unique in their nature. The heterogeneity in mineral composition and physical form, and the variability imposed by an ore body, complicates precise product characterization. There are robust conventions emerging for defining Ores and Ore Concentrates composition more accurately, such as listing constituents in mineralogical terms, rather than chemical (i.e. molecular) terms. More information can be found in the Euromines/ICMM guidance (2009).

6.6.5 Developing categories for complex inorganic UVCB substances

Complex inorganic substances are complex materials containing varying amounts of metals, metal compounds and/or metal minerals. In addition to variability in composition, there is often a large heterogeneity of physical forms. These aspects should be taken into consideration in order to ensure the appropriate assessment of the UVCB. In view of the range of inorganic UVCBs, this would require a vast amount of testing. In addition, the results of these tests would not be predictive due to the temporal and spatial uniqueness of the UVCB.

In order to address the variability and carry out the most adequate assessment, a grouping approach has been developed, as outlined below. As a first step, this entails collecting information on the characterization of the UVCB and its properties. The next step is the selection of (a) representative UVCB (s). The way in which the representative(s) is (are) selected depends on the purpose of the grouping.

First step: Collection of information

Characterization

For characterising the UVCB, information on the following attributes may be available:

- Information on the origin of the inorganic UVCB and the production process (if applicable)
- Chemical (elemental) composition of the inorganic UVCB, to the extent measurable and/or predictable (for example, expert judgement on the basis of the information relative to the feed material)
- Mineralogical composition or species and crystallographic/mineralogical form in which major (elemental) constituents are present in the UVCB – this may provide a first assumption on how (bio-)available each constituent is in the UVCB
- Typical concentrations and/or concentration ranges in which each constituent is/may be present in the UVCB, cut off concentrations in case of hazard triggering constituents

Properties

Information on the physical-chemical, toxicological and eco toxicological properties of the inorganic UVCB should be gathered. When information is not available for the UVCB as such, then information on its constituents should be used:

- Physical-chemical properties as for example the dustiness, granulometry, density, oxidation, melting point etc.
- (Eco)toxicity reference values
- M-factors and Transformation/Dissolution information
- Bio-elution information

Second step: Selection of a representative for grouping and read-across:

Although testing is technically feasible, the wide variability in composition of some inorganic UVCBs across manufacturers, process steps, process streams and over time makes it almost impossible to select an absolutely representative UVCB.

Depending on the purpose of the grouping (e.g. data gap filling, testing, etc.), a representative UVCB will be selected using part or all of the information listed above. The more detailed information that is available, the more precise the determination of the properties of the UVCB will be, and the more refined the (sub-)grouping. Grouping for the purpose of hazard evaluation and grouping for gap filling differ, and require different levels of qualitative and quantitative information.

If information on the relevant UVCB is not available to the required extent, comparable information on the individual constituents should be used. In this approach, the inorganic UVCB is treated as a mixture of its constituents.

6.7 Metals and inorganic metal compounds

6.7.1 Introduction

The concept of grouping has traditionally been widely used for hazard assessment for certain endpoints and risk assessment of metal substances. The approaches have generally been based on the occurrence and bioavailability of a common metal ion (cation or anion) and reading across within a group to fill the data gaps. Such approaches should only be used when substance-specific data are lacking.

For example, the grouping approach based on the metal ion has been used for the classification and labelling of a number of metal compounds like for example Ni in the EU under the Dangerous Substances Directive (DSD) and the Classification, Labelling and Packaging Regulation (CLP)[50]. Other group entries are based on certain anions of concern such as oxalates and thiocyanates. The grouping approach had also been used for estimating the potency of the effects as well as for their identification. NOAEL(s), NOEC(s) have been read-across from data obtained from water-soluble metal compounds to other water-soluble metal compounds of the same metal, including, in the absence of specific data, to compounds of substantially lower water-solubility, on the assumption of a common metal ion. Examples include EU risk assessments on nickel (Tsakovska and Worth, 2007; ECB, 2008a) and zinc which were endorsed by the OECD (ECB, 2008b, OECD, 2005).

A grouping approach has also been used during the categorization of existing chemicals (metals like Co, V, Zn etc.) on Canada's domestic substances list (Environment Canada, 2003). Substances based on a common metal moiety of concern were grouped for assessment, since metal-containing substances able to release the **same metal ion** can contribute to the cumulative loadings, exposure and effects of that metal ion in the environment (i.e. risk).

Under the EU REACH regulation, grouping approaches have been widely applied by industry to comply with data requirements and for developing testing strategies, for animal welfare and resources reasons. Data-filling approaches (grouping and reading across from

[50] the EU terminology for this type of entry is a "group entry" rather than a category

source to target substance) had to be developed and have been applied among others in the molybdenum, vanadium, cobalt and antimony registrations.

The guidance below is based 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 forums (see e.g., Hart, 2007; Schoeters and Verougstraete, 2007), further complemented by that acquired under the recent EU REACH and CLP Regulations. This guidance is intended to supplement the general guidance in the previous chapters with issues specific to metals and inorganic compounds. It includes some metal-specific examples.

6.7.2 Basis and assumptions underlying the grouping of metal compounds

In general, for metals, large data sets are available for most high volume substances, mainly the soluble salts. Thus, for a wide range of compounds of a given metal, data can be limited and data availability will play an important role in source selection for data gap filling.

The main assumption underlying the grouping of metal compounds is that toxicological and ecotoxicological 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). It is the bioavailability of the metal ion (or a redox form of this ion) at target sites that, besides the toxicity potency, will determine the occurrence and severity of the effects to be assessed. This is a reasonable assumption for the majority of inorganic compounds and some organic compounds (e.g., metal salts of some organic acids), in the absence of demonstrated relative differences in bioavailability. The selection of the metal compounds for which a grouping approach (and reading-across from members of the group to others for which data-filling is required) is relevant should be done with care (see 6.7.3). It is made complex by the occurrence of metal (compounds) in a wide and heterogeneous range of materials, under different forms (inorganic metal compounds, organic metal salts, organometallic compounds, metals in elemental form, metal-metal compounds, metal-bearing minerals, alloys and complex substances) but also by a number of factors that could alter the assumption of commonality:

- Chemical speciation and valence: 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 valences may result in differences in mechanism of action and a variation in toxicological properties. For example, extreme differences in hazards are seen with Cr^{3+} and Cr^{6+} compounds. In some cases, chemical species may be interconvertible, in other cases there is little interconversion between the species.
- Organometallic compound: Organometallic compounds,e.g. tributyltin, will generally have a different mode of action compared with that of the metal ion(s) 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, if an organometallic compound degrades in a relevant timeframe under environmentally or biologically relevant conditions to its inorganic metal moiety, it can be considered as a source of the inorganic metal moiety. The potential hazard of liberated organic components should also be considered.

- Metals (elemental form, zero valence): 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 (ECB, 2006a), whilst for others it has not (e.g. soluble nickel salts to nickel metal (ECB, 2008b)). Therefore the appropriateness of read-across needs to be evaluated on a case-by-case basis. It must be noted that the transformation from the metallic form to the ionic state is a corrosion process, in which the metal reacts with air (oxygen) and/or water. Corrosion is defined as *"...chemical or electrochemical reaction between a material, usually a metal, and its environment that produces a deterioration of the material and its properties"* (ASM, 1987). This process is substantially different from a pure dissolution process, and is strongly dependent on the electrochemical properties of the metal, the surface properties and the composition of the medium and biological condition when the process occurs inside an organism.
- Metal containing UVCBs and special mixtures like alloys, glass: 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[51], ashes and drosses[52]. 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, in case of chemical inertness); however, a component based approach can be considered (c.f. section 6.6.).
- 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 synthetic amorphous silicas assessed within the OECD Cooperative Chemicals Assessment 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]). Another relevant example is given by the inorganic pigments: most of them being of a spinel[53] character, individual atoms in their crystal lattice structure can be substituted within ranges without altering the chemical inertness of the spinel itself, thus rendering them bio-unavailable. When this range is exceeded, a metal constituent may become readily leachable from the pigment matrix, thus subjecting the pigment of such a composition to be placed in a different chemical category.
- Particle size and surface properties information: Particle size of the substance influences the deposition behaviour in the respiratory tract, rates of dissolution and

[51] Frits are ceramic compositions, that are fused in an oven, quenched to form a glass, and granulated.
[52] A dross is a solid impurity(ies) floating on a molten metal or dispersed in the metal, such as in wrought iron. It forms on the surface of low-melting-point metals such as tin, lead, zinc or aluminium or alloys by oxidation of the metal(s).
[53] Spinel – The spinel structure is a crystallisation pattern for a multitude of minerals. Spinels can be synthesised with the common structural formula [AXB2-X]O4, "A" being a divalent metal and "B" a tri- or tetravalent metal, via sintering, hydrothermal synthesis or the Verneuil process. Since various metals can crystallise in the spinel structure, countless modifications exist. Spinels, especially the synthesised high temperature modifications, are characterised amongst others by their hardness (Mohrs scale ca. 8) and overall chemical inertness, which makes it of particular interest as pigments, gemstones and in ceramics.

corrosion which may result in significant changes to the toxicological profile. Based on particle size distribution data and surface properties, trends in deposition and potency of effects can be assessed for locally acting substances, for example in the case of Ni sulfide particles. When these have negative surface charge these are taken up by epithelial cells much more effectively than if these are positively charged, which can affect their potency. If there is evidence that the surface properties and particle size influence the bioavailability significantly and therefore impact the severity of the effects of the compound to be assessed, this should be taken into account in a WoE approach considering all available information (e.g. toxicokinetics. Particle size can also influence the extent of solubility in aqueous and biological media. Note this section is not intended to address nanomaterials which are discussed briefly in Section 6.9.

- 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. 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 significantly influence the effects of the compound to be assessed and alter the assumption of commonality, this should be taken into account in the evaluation.

6.7.3 Grouping of metals: in practice

Water solubility of the metal compounds has previously been used as the starting point for establishing a group, as it was estimated to provide a first indication of the relative availability of the metal ion in aqueous media. The most simplistic approach to hazard evaluation assumes that the metal ion availability will normally be reduced with decreasing water-solubility and consequently present a reduced bioavailability. In this cautious approach a specific metal containing compound will be evaluated as showing the same hazards as the most water-soluble compounds of that metal in the group. The approach may be protective for hazard assessment in aqueous systems but not for sediments and soils.

This approach can be refined for by building subgroups based on water solubility, when data are available on toxicity trends with water solubility or when data-rich reference compounds exist for several water solubility categories. For example, for inorganic nickel compounds a number of sub-groups had been suggested, reflecting different ranges of aqueous solubility (Hart, 2007). As another 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 also, to some extent recognised in the evaluation method used for the ***environmental hazard classification*** of metals and metal compounds, where the relevant hazard categories can be evaluated using a transformation/dissolution protocol (OECD, 2001b), which compares the amount of released metal ion from the metal compound within defined timeframes, with classification thresholds as a function of pH between 6 and 9 (a pH range which may typically be encountered in the environment).

1. Grouping for human health endpoints:

For ***human health endpoints***, recent work on e.g. nickel compounds has shown that the use of water solubility alone may be either too precautionary or not conservative enough (Henderson *et al.,* 2012). Using water solubility as surrogate for bioavailability for human health is associated with uncertainties that should be acknowledged: metals do not dissolve in water; in fact, these react with water or oxygen in a process more appropriately described as "corrosion". Water solubility in distilled water is driven by the solubility product of the anionic and cationic moieties. Such water solubility data do not reflect the influence of different pH or redox conditions, or the presence of various ligands. Chemical speciation in physiological fluids present in various body compartments may render use of water-solubility data less preferable. The effects of particulates are not addressed here.

More refined weight-of-evidence approaches using additional physical-chemical property information such as bioaccessibility in synthetic biological fluids have been published for nickel (Henderson *et al*., 2012), and have been used for registration under REACH for a number of metals. Bioaccessibility test results may indicate the extent of metal ion release in vitro in fluids mimicking relevant physiological fluids and are considered to provide a better basis for estimation of in vivo bioavailability and hence also for uptake and systemic toxicity.

Therefore the proposed approach for human health endpoints is explained hereunder and illustrated with examples. This approach has been presented at a workshop arranged by the metals industry, for the EU Commission, ECHA and the EU Member States on October 1 2012 (Eurometaux, 2012). A specific case on employment of the approach on a proposed cobalt grouping has been discussed and approved in the EU Member States Committee under REACH in June-September 2013.

Step 0: Determine if the metal compound is already a member of a group or existing category.

Step 1: Generate metal release data in an appropriate bioaccessibility testing set up (using appropriate fluids related to the route of exposure considered) for source and target substances. A preliminary grouping can be done based on these data.

Step 2: Consider the bioaccessibility data in WoE approach with data on physical-chemical properties (e.g., water solubility, degree of dissociation of the metal-containing compound/mineral, particle size and structure), knowledge on mode and if possible mechanism of action (in particular for local effects), and factors like presence of counter-ions. Incorporate existing in vivo data and as appropriate new targeted *in vivo* toxicity and/or toxicokinetic testing in order to verify that the bioaccessibility data correlate with the toxicity endpoint(s) considered and to generate a reference range. In relation to establishment of a chemical category: Pay particular attention to the need to robustly define the borders of the category (the most comprehensive test data are required for the substances in each end of the category).

Step 3: Assess the most appropriate grouping of substances and identify the source substance for each target substance based on the weight-of-evidence approach described above.

Step 4: Use the new paradigm to read-across toxicological data from source substances to target substances based on the weight-of-evidence approach described above.

Illustrative Example:

Nickel: In the case of nickel, bioaccessibility data (Ni^{2+} release) in synthetic gastric fluid (2 hours) and intestinal fluid (24 hours) were gathered from a dozen different nickel compounds (step 1). Acute oral toxicity data (LD_{50}) available for a subset of these compounds was used to verify that grouping based on relative Ni release in gastric fluid was appropriate to read across hazards associated with systemic (oral) exposure to nickel (Henderson *et al*., 2012a,b) (step 2). In step 3, the preliminary grouping based on bioaccessibility was confirmed by the *in vivo* verification and in step 4, read-across for acute oral toxicity hazard classifications was performed (http://www.reach-metals.eu/index.php?option=com_content&task=view&id=44&Itemid=41).

2. *Grouping for environmental endpoints:*

A grouping approach is followed based on all information available for inorganic substances. The grouping is based on the assumption that properties are likely to be similar or follow a similar pattern as a result of the presence of the free metal ion. For most metal-containing substances, it is indeed the potentially bioavailable metal ion that is liberated (in greater or lesser amounts) upon contact with water that normally is the moiety of toxicological concern.

This assumption can be considered valid when i) minor differences in solubility among the metal compounds do not result in significant differences in ecotoxicity, ii) ecotoxicity is only affected by the metal ion and not, or far less, by the counter ions, and iii) there are no important differences in speciation of the metal from the different metal substances within the proposed group in the environment after emissions.

Illustrative example:

Cobalt and Cobalt compounds: the understanding of the physical-chemical and ecotoxicological data for metal and counter-ions is essential to both understand the environmental fate and toxicological characteristics of the Co compounds and to provide support for grouping approach using results obtained in tests conducted with soluble cobalt salts (e.g., cobalt dichloride) for all inorganic Co compounds that were tested. In most cases environmental fate and effects data developed with the free counter-acid, or a simple salt that would readily dissociate (e.g. the sodium salt), can serve as surrogate data for the anionic component of each cobalt salt. Similarly, data for the metal ions can be represented by fate and toxicity data generated with simple metal salts (e.g. chloride or nitrate salts). For example, the potential hazards associated with cobalt acetate can be estimated through the evaluation of the cobalt free ion, tested as cobalt dichloride, and the acetate moiety, tested as sodium acetate. Thus, data for each individual counter-ion (tested as the free counter-acid or Na, K or Ca salt) and the individual metal (tested as the metal chloride or other simple metal salt) can be used to "read-across" to characterise the hazard of a cobalt compound.. For the cobalt compounds tested the chronic EC_{10} values were statistically similar to one another, suggesting that the free cobalt ion diominates the chronic toxicity of all of the substances included in the proposed grouping. This approach was applicable for all relevant environmental endpoints (CDI, 2009; http://webnet.oecd.org/Hpv/UI/handler.axd?id=a67489b5-6fa0-40ca-8938-c0bbce002d24)

6.7.4 Need to verify the assumptions and further refinements

When applying grouping approaches, one should verify whether the approach is sufficiently robust for the assessment purpose. As indicated above, confidence in the application of grouping approaches and performing reading -across from one metal compound (source) to another with the same metal ion (target) is strongest for substances demonstrating similar toxicological effects and *bioavailability*.

Note1: for environment, effect data sets are preferably to be assessed at equal bioavailability level. Several models have been developed to correct for bioavailability of metals in aquatic media, soil and sediments (e.g. Biotic Ligand Models (BLM) for metals in aquatic systems (Paquin *et al.,* (2002) and Niyogi and Wood (2004), and ECHA REACH guidance (ECHA, 2008).

Note 2: for human health, in order to e.g. verify the bioaccessibility based grouping, following options could be considered:

- In vitro data:

In vitro information on tests such as *in vitro* test for skin irritation could be used to verify the grouping and read-across based on bioaccessibility in synthetic sweat[54].
For respiratory local effects after inhalation, the physical-chemical properties of the metal-containing particles (such as particle size distribution and surface area[55]) can significantly influence the local toxicity and therefore, the assessment of bioavailability of the metal ion could be complemented with relevant and reliable *in vitro* and/or *in vivo* information. For example, *in vitro* measurements of relative uptake of particles by lung epithelial cells or ability of particles to trigger reactive oxygen species release from macrophages may be informative.

- 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 that skin (or eye) 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 EPA, 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 required or considered necessary. However, where such information is not available, information could be collected for representative members of the category. It is obvious that assessments as those described here will require thorough case by case and WoE based expert judgement.

[54] No OECD guideline is currently available.
[55] The current OECD Guideline is OECD TG 110 from 1981

6.8 General guidance for other compounds (e.g. ionisable 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 *et al*., 2007). One example is the Methanolates category assessed under the OECD Cooperative Chemicals Assessment Programme (www.oecd.org/env/existingchemicals/data). This consists of 17 potassium and sodium methanolate and both react rapidly in water to form the corresponding hydroxide.

When comparing acids their salts and multiprotic molecules, differences arising from pH effects should be considered (Caley *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 Cooperative Chemicals Assessment Programme

(www.oecd.org/env/existingchemicals/data). For these categories, dermal and irritation studies are considered separately for the acid and salts.

For the Gluconates category assessed under the OECD Cooperative Chemicals Assessment Programme (www.oecd.org/env/existingchemicals/data) , it was found that for categories including ionisable compounds, the effect of the counter-ion needs to be considered (Caley *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.

6.9 Initial considerations applicable to manufactured nanomaterials

Nanomaterials[56] are a subset of chemicals that are distinguished by their size (in general between 1 and-100 nm. Most materials described as nanomaterials are solids though there are also liquid nanoparticles such as in some emulsions. Principles and guidance for grouping nanomaterials for the purpose of assessing their toxicological, ecotoxicological and fate properties, are under development. Fundamental research is currently devoted to identifying and characterizing exposure paths, bioavailability and bioactivity of nanomaterial both *in vitro* and *in vivo* and the results of this research will provide information that may be useful for selection of appropriate analogues and categories.

Nanomaterials share properties associated with both solutes and separate particles phases, features that complicates the risk assessment of nanomaterials include i) their measurements and characterization in environmental and biological matrices for understanding their fate, transport, and potential impact, and ii) their preparation and testing procedures for the assessment of their bioavailability and effects on organisms (Burello and Worth, 2011; Alvarez *et al.*, 2009; OECD 2012f and 2012g), iii) the many physical-chemical characteristics (e.g. size, coating, shape, surface characteristics, solubility) that can influence fate, behaviour,

[56] A definition is given at http://ec.europa.eu/environment/chemicals/nanotech/faq/definition_en.htm

kinetics and toxicity, and iv) the potentially constantly changing physical-chemical characteristics during the life-cycle of a material. Among initial considerations for characterizing nanomaterials, properties such as structure, size, shape, surface area, surface modification, surface reactivity and electronic properties, agglomeration state and water solubility are certainly relevant to predict their bioavailability, and the mechanism of action potentially leading to effects on organisms and behaviour in the environment. Any guidance likely to be developed in the future will be based on experimental data as well as modelled data and predictions. There are several initiatives in OECD countries to generate good quality data on representatives types of nanomaterials (e.g. OECD sponsorship programme), there are also numerous projects in OECD member countries aiming at developing computational approaches to predict properties of nanomaterials, e.g. oxidative stress potential of oxide nanoparticles (Gallegos Saliner *et al*., 2009, Burello and Worth, 2011). At present, it seems premature to develop guidance on grouping specifically for nanomaterials. Nevertheless, research efforts will pave the way for common approaches and frameworks to grouping nanomaterials for purpose of hazard assessment in the future. In addition, expand further on why certain properties tend to elicit certain effects in vitro or in vivo and where opportunities may exist to group nanomaterials together to rationalize testing. Section 6.9 will be amended as accepted principles for grouping and read-across of nanomaterials arise from these activities.

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 analogue or chemical category approach in e.g. Chapter 1 of a SIDS Initial Assessment Report and in the SIDS Initial Assessment Profile in a summarised form or Chemical Safety Report. Illustrative examples are provided in the ECETOC technical report (ECETOC, 2012) and its complementary manuscript (Patlewicz *et al*., 2013). ECHA have also recently published an illustrative example (ECHA, 2013a)

Importantly, the documentation on the category definition and justification, as indicated in section 2.5 and chapter 3, Robustness of a chemical category, needs to be provided.

For chemical categories, the assessment report should address all members of the chemical category and be accompanied by robust study summaries of the key studies for all relevant endpoints (physical chemical properties, environmental fate and pathways, ecotoxicity, toxicity) for each member of the category. In the case where analogue data is used in the assessment from a source chemical that is not a formal category member, the robust study summary(ies) for the endpoints being read-across need to be provided in the dossier of the target chemical. Alternatively, a separate dossier can be provided for the source chemical, especially if several endpoints are being read-across.

Experience in the OECD HPV Chemical Programme has shown that for a simple analogue approach, 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) , as indicated above.

Furthermore, when developing an analogue or chemical category approach with IUCLID 5 or any other similar software having implemented the OECD harmonised templates (OECD, 2006e), 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 End User Manual (EC, 2007) and User Manual (ECHA, 2013b).

7.1 Reporting Format for analogue approach

1. Hypothesis for the analogue approach

Provide the chemical descriptor common identifiers (including CAS number) and structures as far as possible of the source and target substances. This information is critical so relationships between the target and the source substance(s) can be clearly identified.

Describe the molecular structure a substance must have to be suitable as a source substance. All functional groups need to be identified. Provide the hypothesis for why the read-across can be performed. The structural and mechanistic similarities between the target

chemical and the analogue(s) should be identified (e.g., functional group, moiety of concern, carbon chain length, common metabolic or degradation pathways and products). If there is a mechanistic reasoning to the read-across, describe the foreseen mode of action or adverse outcome pathway for source and target chemicals and if relevant describe the influence of the mode of administration of the source chemical (oral, dermal, inhalation) and its relevance in relation to the physical form of the target chemical, (see section 2.4 and 2.5 for more guidance).

List the endpoints for which the analogue approach is applied; for other endpoints, relevant information should be available. Depending on functional group(s), reactivity, mechanism of action, the analogue approach may apply for some endpoints only (e.g. acute effects only) and this should be specified and justified.

If the hypothesis is based on the assumption that the chemical or biological conversion of target and source substances results in exposure to the same toxicants, and subsequently to the same effects, the consideration outlined in the end of section 2.5 should be taken into account.

2. Source chemical(s)

Describe the source chemical(s) as comprehensively as possible. Provide CAS numbers, names and chemical structures of the source substance (s) (see for more information section 2.3.1).

3. Purity / Impurities

Provide purity/impurity profiles for the target and source substances, including the likely impact on the relevant endpoints. It should be discussed which influence these impurities are thought to have on physical-chemical parameters, fate and (eco)toxicology, and hence on the read-across. See Section 3.2.3.1.5.

4. Analogue approach justification

Based on available experimental data, including basic physical-chemical properties, available toxicity and ecotoxicity data from the source and the target substances summarise how these results verify that the read-across is justified. It should be explained why the read-across between the source chemical to the target chemical is actually justified, and why the analogue(s) used are adequate and provide sufficiently robust information to characterize the hazard endpoint(s) considered. The data should also show that functional groups not common to source and target substances 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, reliability of the experimental data, variability of experimental results between source and target substances, the quality of the data estimated by external computational approaches 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. Strategy used to fill the data gaps

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
- In the matrix, data gaps filled by read-across from analogues should be highlighted

If experimental results are available, the key study results should be shown in the Data Matrix. Table 11 below is an example of how to develop a data matrix for an analogue approach. All source substances relevant to the target substance should be included, as should all available experimental data. In addition to providing the data, it is also important to describe similarities and/or trends in data that would support an analogue approach. Important to note is that similarities and/or trends may not be observed for all types of endpoints. This discussion can be included in the data matrix, as shown, or, alternatively, described in the analogue justification on an endpoint basis.

Table 11. Data matrix, analogue approach

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 (logKow)					
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. The weight of evidence assessment for skin sensitisation potential for 4-isopropylaniline (OECD, 2014b (in prep.) gives an example on how the analogue approach could be reported.

7.2 Reporting format for chemical categories

The logical framework provided in section 2.5 and chapter 3 should be applied and structured as follows:

1. Category Definition

Provide a summary of the common features of the category members; describe the boundaries (e.g., in number of carbon atoms); physical-chemical properties, if applicable (e.g., boiling point); allowed variations in chemical structure; and if known, any restrictions (e.g., variations that would change the effects of a substance significantly compared to the other substances in the category).

2. Category Members

For each category member, provide the CAS number, name and chemical structure. Describe all the category members as comprehensively as possible. Section 2.3.2 gives information on category membership and Section 3.2.3 details the elements for a read across justification including the chemical identity and compositional aspects.

The structural and functional elements and the relationship between the various category members need to be stated in a clear and unambiguous manner. Table 12 gives examples of how to build a matrix of the category members and to present the elements that change or stay constant within the category, and to provide the most representative structures.

Table 12. Examples for structural matrix

Example	Carbon number	Branching type	Functional group	Position of functional group	Most Representative structure(s)
Substance 1	C9	Linear	e.g., terminal OH	Alpha	Structure 1
Substance 2	C11	Branched		Beta	Structure 2
Substance 3	C13	Iso [two methyl groups on the backbone carbon chain]			Structure 3

Structural elements will be specific to a category and could be such items as

- Salts
- Carbon number of chain
- Degree and nature of branching or occurrence of double bonds, functional groups, aromatics, cycles, hetero-cycles)
- Moiety

- Valency
- Positioning of the common functional element

Any other aspects that may be important to the development of the category, for example boiling point for hydrocarbon streams, should be included. The objective is to build an overall picture of the validity domain of the proposed category by defining the relationships between its members and setting the boundaries in structure and its chemical properties. Analytical data of the chemical structures of category members may be useful to demonstrate how the structural properties change over the category.

Compositional Information

As described in Section 3.2.3.1.2, while structure is critical for some categories, composition within the category members may be the critical factor for others. This is especially true for multi-constituent substance and UVCB categories and examples are given of categories based upon component analyses (See table 2).

An approach is described for reporting upon and characterising category composition (sections 3.2.3.1.4 and example table for reporting upon category composition is also given. Such a table should be considered to give the maximum clarity to the bounds of the category and the proposed read across.

Section 2.3.2 gives information on category membership and Section 3.2.3 details the elements for a read across justification including the chemical identity and compositional aspects.

Purity / Impurities

As described in Section 3.2.3.1.5 impurities are important to consider and may relate to the source and manufacturing route of the substance under consideration. Consequently it is of value to 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 may have on physical-chemical parameters, fate and (eco)toxicology as well as toxicological properties, and hence on the read-across. See Chapter 5 for further information on building a category. An example of an approach for UVCBs is provided in Chapter 6, Section 6.6.

Physical-chemical properties

As described in Section 3.2.3.2 physical-chemical properties are a critical determinant to the environmental and health properties of a substance affecting bioavailability, environmental fate, and thus the (eco)toxicity of a chemical. Similarly, these may also impact the toxicology of a substance. Consequently physical-chemical properties across a category should be elaborated as part of its basic properties and reported in a table. A plot of a trend is usually very helpful to give clarity. As described previously, read across between category members for physical-chemical properties is not usually a good practice as the data are required to demonstrate the trends, or lack of, in the category and data can be generated through different techniques.

3. Category Hypothesis

Describe the molecular structure a substance must have to be included in the category. Provide the 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, i.e. oral, dermal, inhalation (see for more information section 2.4).

If the hypothesis is based on the assumption that the chemical or biological conversion of target and source substances results in exposure to the same toxicants, and subsequently to the same effects, the consideration outlined in the end of section 2.5 should be taken into account. Many of the other potential bases for a category hypothesis are discussed in Chapter 3 and examples are given.

4. Category justification

Based on available experimental data (including appropriate physical-chemical data and additional test results and molecular descriptor or profiler values that might have been generated for the assessment of this category) and knowledge of metabolism or mode and/or mechanism of action or adverse outcome pathway, summarise how these results for each included substance in the category verify that the category is robust. This should include an indication of the trend(s) for each endpoint, if such trend exist, and what explains the trend observed (e.g. incremental structural changes) and whether such trend applies to the whole category or whether breakpoints/thresholds are to be expected. Alternatively, when no trend appears clearly across the category, the strategy applied for the read-across should be stated (e.g. read-across from the closest analogue, average from several analogues, or worst case scenario). 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 category and the read-across.

The existence of sub-categories and the rationale for sub-categorising (e.g. existence of thresholds in physical-chemical properties impacting solubility or bioavailability of category members and thus hazards) should be provided.

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).

5. 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 substances 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.

6. 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). Specify carefully, and as comprehensibly as possible, why it is justified to employ the proposed category approach for the endpoints suggested. Link an explanation that covers all proposed category members and that justifies that the structurally related properties used to create the category are plausibly related to the endpoints suggested to be covered by the proposed category approach.

7. Strategy used to fill the data gaps

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) to add clarity to trends in any relevant properties. For example, the ordering of the members should reflect a trend or progression within the category.

The read-across strategy used should be explained at this stage: for each endpoint and chemical, the availability of experimental data needs to be indicated. In the case where no data are available, the method for read-across needs to be specified (closest category member? worst case scenario?), with justification for the selected strategy. In case of sub-categories, these should be easily identifiable in the data matrix table.

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[57]
- (Q)SAR

In the matrix, data gaps filled by read-across from category members should be highlighted.

If experimental results are available, the key study results should be shown in the Data Matrix. Table 13 below is an example of how to develop a data matrix for a chemical category. All potential category members should be included, as should all available experimental data. In addition to providing the data, it is also important to describe trends in data that would justify a category approach. Important to note is that trends may not be observed for all types of endpoints. The category justification can be included in the data matrix.

[57] 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 pick list for the data element "study result type". Instead the item "read-across based on grouping of substances (category approach)"could be used.

It is useful to indicate which analogues are used for read-across, to provide detailed information on the data used for data gap filling (result, test type, study design), and to specify the technique used for estimating the value of the target substance (e.g., read-across, trend or (Q)SAR) as it allows proper evaluation of the category justification and predictions derived from it.

It might be useful to build more than one matrices in a report. One can simply indicate the presence of experimental data with sufficient reliability for read-across. Others could include an overview of all collected data, data gaps that remained and that were filled by read –across. The report should be readable also as stand-alone document.

More detailed discussion of how data gaps are filled for individual endpoints and individual category members (e.g. read-across, trend analysis, (Q)SAR) as well as the rationales for the chosen technique of filling the data gaps should be provided in the corresponding sections of the assessment report.

For UVCB substances 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.

Table13. 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 (logKow)					
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					
…					

* For data-rich substances, the matrix could become very large, and could therefore be broken down into groups of endpoints.

8 REFERENCES

Aptula, A.O., G. Patlewicz, D.W. Roberts and T.W. Schultz (2006), Non-enzymatic glutathione reactivity and in vitro toxicity: a non-animal approach to skin sensitization, *Toxicology In Vitro*, Vol. 20, No. 2, pp. 239-247.

ASM (1987), Corrosion, Ninth Edition, Volume 13, ASM International.

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

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

Beischlag, T.V., J.L. Morales, B.D. Hollingshead, and G.H. Perdew (2008), The Aryl Hydrocarbon Receptor Complex and the Control of Gene Expression. *Critical Reviews™ in Eukaryotic Gene Expression*, Vol. 18, No. 3, pp. 207–250.

Benigni, R., L. Conti, R. Crebelli, A. Rodomonte and M.R. Vari (2005), Simple and alpha, beta-unsaturated aldehydes: correct prediction of genotoxic activity through structure-activity relationship models, *Environmental and Molecular Mutagenesis*, Vol. 46, pp. 268-280.

Bishop P.L., J.R. Manuppello, C.E. Willett, and J.T. Sandler (2012), Animal use and lessons learned in the U.S. High Production Volume Chemicals Challenge Program, *Environmental Health Perspect*, Vol. 120, No. 12, pp. 1631-1639.

Burello, E, and A. Worth, (2011), Computational nanotoxicology: Predicting toxicity of nanoparticles, *Nature Nanotechnology*, Vol. 6, 138–139

Caley, J., T. Henry, D. Merckel, M. Sonowane and A. Benson (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.

Canada, E. (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.

Canada (2010a), Phenol, (1,1-dimethylethyl)-4-methoxy-(Butylated hydroxyanisole), Screening Assessment for the Challenge.

Canada (2010b), Decanedioic acid, bis(1,2,2,6,6-pentamethyl-4-piperidinyl)-ester. Screening Assessment for the Challenge.

Canada (2010c), Final Second Priority Substances List Assessment Report on Three Aluminum Salts.

Canada (2010d), Methylium, [4-(dimethylamino)phenyl]bis[4-(ethylamino)-3-methylphenyl]-, acetate.

Canada (2013), http://www.chemicalsubstanceschimiques.gc.ca/plan/index-eng.php

CDI (2009), Development of Protective Water Quality Concentrations for Cobalt - Data in preparation. Aquatic Research Program. The Cobalt Development Institute. Guildford, Surrey, United Kingdom

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

CESIO (2003), CESIO recommendation for the classification and labelling of surfactants as „dangerous for the environment".

Cefic (2012), Experts Workshop on Read-Across Assessment organised by ECHA with the active support from Cefic-LRI (October 3, 2012)Use of read-across for Chemical Safety Assessment under REACH.

Clemedson C., M. Nordin-Andersson, H.F. Bjerregaard, J. Clausen, A. Forsby, H. Gustafsson, U. Hansson, B. Isomaa, C. Jørgensen, A. Kolman, N. Kotova, G. Krause, U. Kristen, K. Kurppa, L. Romer, and E. Scheers (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,* Vol. 30, pp. 303-321.

Clemens, J.H., M.R. van den Heuvel, J.J. Stegeman, D.G. Dixon, and N.C. Bols (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*, Vol. 14, pp. 280.

Comber, M. and B. Simpson (2007), Grouping of Petroleum substances. In: A Compendium of Case Studies that helped to shape the REACH Guidance on Chemical Categories and Read Across. EUR report no 22481 EN.

Concawe (1996), Environmental risk assessment of petroleum substances: the hydrocarbon block method. Report no. 96/52

Concawe (2012), Hazard classification and labelling of petroleum substances in the European Economic Area – Report no. 8/12

Concawe (2012), REACH – Analytical Characterisation of Petroleum UVCB Substances. Concawe document number 7/12. Concawe, Brussels, Belgium.

Council Reg. (EEC) No. 793/93 of 23 March 1993 on the evaluation and control of the risks of existing substances.

http://ec.europa.eu/environment/chemicals/exist_subst/pdf/793_93_ec.pdf

Creton, S., S.A. Saghir, M.J. Bartels, R. Billington, J.S. Bus, W. Davies, M.P. Dent, G.M. Hawksworth, S. Parry and K.Z. Travis (2012), Use of toxicokinetics to support chemical evaluation: Informing high dose selection and study interpretation, *Regulatory Toxicology and Pharmacology*, Vol. 62, No. 2, pp. 241-247.

Cruywagen, J.J. (2000), Protonation, oligomerization, and condensation reactions of vanadate(V), molybdate(VI), and tungstate(VI), *Advances in Inorganic. Chemistry*, Vol. **49**, pp. 127-182

Cruywagen, J.J., A. G. Draaier, J.B.B. Heyns and E.A. Rohwer (2002), Molybdenum(VI) equilibria in different ionic media. Formation constants and thermodynamic quantities, *Inorganic Chemistry Acta*, Vol. 331, pp. 332-329.

Danish (Q)SAR Database: http://qsar.food.dtu.dk/

Derek (LHASA Ltd, UK): http://www.lhasalimited.org/products/derek-nexus.htm

Diderich, R. (2010), Tools for category Formation and Read-Across: Overview of the OECD (Q)SAR Application Toolbox. In Silico Toxicology, Principles and Applications- Cronin& Madden, The Royal Society of Chemistry. Chapter 16(7).

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

EC (2007), IUCLID 5 End User Manual.

EC DG JOINT RESEARCH CENTRE (2008) Institute for Health and Consumer Protection QSAR Prediction Reporting Format (QPRF) (version 1.2) http://ihcp.jrc.ec.europa.eu/our_labs/computational_toxicology/qsar_tools/qrf/QPRF_version_1.1.doc

EC (2008), Regulation (EC) No 1272/2008 of the European Parliament and of the Council of 16 December 2008 on classification, labelling and packaging of substances and mixtures) amending and repealing the Dangerous Substances Directive 67/548/EE and the Dangerous Preparations Directive 1999/45/EC, and amending Regulation (EC) No 1907/2006 http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CONSLEG:2008R1272:20110419:EN:PDF

ECB (2005), Scoping study on the development of a technical guidance document on information requirements on intrinsic properties of substances (RIP 3.3-1), CEFIC, DK-EPA, Environmental Agency of Wales and England, ECETOC, INERIS, KemI and TNO. European Chemicals Bureau, Joint Research Centre, European Commission, Ispra, Italy.

ECB (2006a), European Union Risk Assessment Report on, Cadmium oxide CAS-No.: 1306-19-0 and Cadmium metal CAS-No.: 7440-43-9. Unpublished draft, European Chemicals Bureau.

ECB (2008a), European Union Risk Assessment Reports on Nickel and Nickel Compounds. ECB: 2008., European Chemicals Bureau.

ECB (2008b), European Union Risk Assessment Report Zinc metal, Vol. 42. Final report, (S.J. Munn *et al*., eds.), European Chemicals Bureau**:** 812.

ECB (2008c), European Union Risk Assessment Report Pitch, coal tar, high-temp ECB: 2008, European Chemicals Bureau.

ECETOC (2010), http://www.ecetoc.org/tra

ECETOC (2012), Technical Report 116, Category approaches, read-across, (Q)SAR.

ECHA (2008), Guidance on information requirements and chemical safety assessment Appendix R.7.13-2: Environmental risk assessment for metals and metal compounds.

ECHA (2012a), An Introduction to the Assessment of Read-Across in ECHA. http://echa.europa.eu/documents/10162/5649897/ws_raa_20121003_background_paper_an_introduction_to_the_assessment_of_read-across_in_echa_en.pdf

ECHA (2012b), ECHA guidance for identification and naming of substances under REACH and CLP.

http://echa.europa.eu/documents/10162/13643/substance_id_en.pdf

ECHA (2013a), Grouping of substances and read-across approach. Parts 1 and 2. http://echa.europa.eu/web/guest/support/grouping-of-substances-and-read-across . Last accessed 23 May 2013

ECHA (2013b), IUCLID End User Manual http://iuclid.eu/index.php?fuseaction=home.documentation&type=public#usermanual

EDTA
(https://www.google.be/search?q=ETAD%2C+2001&oq=ETAD%2C+2001&aqs=chrome.0.57.662j0&sourceid=chrome&ie=UTF)

EEC (1993), from http://webnet.oecd.org/hpv/ui/handler.axd?id=fb782359-0b8f-40e1-a951-3d28db606af0

EPA HPV Chemical Hazard Characterizations

Environment Canada. (2010), Domestic Substances List
http://www.ec.gc.ca/lcpe-cepa/default.asp?lang=En&n=C6C230D5-1

EPA (2000), Test plan for cinnamyl derivatives. Prepared by the Flavor and Fragrance High Production Volume Consortia..

EPA (2001a), High Production Volume (HPV) Challeng. Test plan for dicarboxylic acid category. http://www.epa.gov/chemrtk/pubs/summaries/thiobpad/c13379tc.htm

EPA (2002a), High Production Volume (HPV) Challenge. Robust Summaries & Test Plans: Sulfosuccinates. http://www.epa.gov/hpv/pubs/summaries/terprial/c12965tc.htm

EPA (2002b) High Production Volume (HPV) Challenge. Robust Summaries & Test Plans: Benzyl Derivatives.
http://www.epa.gov/hpv/pubs/summaries/benzylde/c13450tc.htm

EPA (2003a), High Production Volume (HPV) Challenge. Robust Summaries & Test Plans: Mononitroanilines Category.
http://www.epa.gov/hpv/pubs/summaries/mntranlc/c14391tc.htm

EPA (2003b), High Production Volume (HPV) Challenge. Robust Summaries & Test Plans: Fuel Oils Category.
http://www.epa.gov/hpv/pubs/summaries/fueloils/c13435tc.htm

EPA (2004a), High Production Volume (HPV) Challenge. Robust Summaries & Test Plans: Pyridine and Pyridine Derivatives Category.
http://www.epa.gov/hpv/pubs/summaries/pyriderv/c14925tc.htm

EPA (2004b), High Production Volume (HPV) Challenge Robust Summaries & Test Plans: Terpenoid Primary Alcohols and Related Esters Category.

http://www.epa.gov/hpv/pubs/summaries/terprial/c12965tc.htm

EPA (2004c), High Production Volume (HPV) Challenge. Robust Summaries & Test Plans: Phosphoric Acid Derivatives.
http://www.epa.gov/hpv/pubs/summaries/phsacdde/c13356tc.htm

EPA (2005), High Production Volume (HPV) Challenge. Robust Summaries & Test Plans: Cinnamyl Derivatives.
http://www.epa.gov/hpv/pubs/summaries/cinna/c12912tc.htm

EPA (2007a), High Production Volume (HPV) Challenge. Robust Summary: Trimellitate Category

EPA (2007b), High Production Volume (HPV) Challenge. Robust Summaries & Test Plans: Cyclohexyl Derivatives Category. (http://www.epa.gov/hpv/pubs/summaries/cyclodrv/c14686tc.htm

EPA (2010), High Production Volume (HPV) Challenge. Robust Summaries & Test Plans: Thiobis Propanoic Acid Derivatives.
http://www.epa.gov/chemrtk/hpvis/hazchar/Category_Thiobis%20Propanoic%20Acid%20Derivatives_June%202010.pdf)

EPA (2013), High Production Volume Challenge.

EPA, U. (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.

EPA, U. (2009), Methodology for Hazard-Based Prioritization under ChAMP. Office of Pollution Prevention and Toxics, Washington DC.

EPA EPISuite US EPA. (2012), Estimation Programs Interface Suite™ for Microsoft® Windows, v 4.11. United States Environmental Protection Agency, Washington, DC, USA. http://www.epa.gov/opptintr/exposure/pubs/episuite.htm

ETAD (2000), Ecological and Toxicological Association of Dyes and organic Pigments Manufacturers. Guidance on hazard labelling of reactive dyes. Journal Society of Dyers and Colourists, Vol. 116:251-2.

EURL ECVAM, E. (2013), European Union Reference Laboratory for alternatives to animal testing, from http://ihcp.jrc.ec.europa.eu/our_labs/eurl-ecvam

EURL (2013), ECVAM Recommendation on the 3T3 Neutral Red Uptake (3T3 NRU) Citotoxicity Assay for the identification of Substances not Requiring Classification for Acute Oral Toxicity, EUR 25946 EN, 2013.

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

Eurometaux (2012), Short report workshop read-across metals_01102012.

Euromines/ICMM (2009), Ores and Concentrates: An industry approach to EU Hazard Classification, with Arcadis and Arche, London, UK http://www.icmm.com/search-results?sortField=sort_rank&query=ores+and+concentrates.

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

Gallegos Saliner, A., E. Burello, and A. Worth (2009), Review of computational approaches for predicting the physicochemical and biological properties of nanoparticles, EUR 23974 EN.

Greenwood, N. N., A. Earnshaw and (Eds.) (1997), Chromium, Molybdenum and Tungsten, Chemistry of the Elements. 2nd Ed., chap. 23: 1002-1039

Hanway, R.H. and P.F. Evans (2000), Read-across of toxicological data in the notification of new chemicals, *Toxicology Letters,* Vol. 116, No. 1, pp. 61.

Hart, J., G.D. Veith (2007), Applying Chemical Categories to Classification & Labelling: A Case Study of Volatile Aliphatic Ethers, Danish Environmental Protection Agency.

Hart, J. (2007), Nickel compounds – a category approach for metals in EU Legislation. A report to the Danish Environmental Protection Agency.

Henderson, R.G., D. Cappellini, S.K. Seilkop, H.K. Bates and A.R. Oller (2012), Oral Bioaccessibility Testing and Read-Across Hazard Assessment of Nickel Compounds, *Regulatory Toxicology and Pharmacology*, Vol. **63, No.** 1, pp. 20-28.

Hoffman, G.M. (2008), Baseline Gasoline Vapor Condensate: A Two-Generation Whole-Body Inhalation Reproductive Toxicity Study in Rats. Huntingdon Life Sciences, Inc. Study No. 00-4207; sponsored by the American Petroleum Institute

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 http://www.oecd.org/env/ehs/oecdquantitativestructure-activityrelationshipsprojectqsars.htm, O. Q. S.-A. R. P.

IPCS (2004), Risk assessment terminology. Geneva, World Health Organization, International Programme on Chemical Safety

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.

IPIECA (2010), Guidance on the application of Globally Harmonized System (GHS) criteria to petroleum substances. IPIECA, Version 1. http://www.ipieca.org/publication/guidance-application-globally-harmonized-system-ghs-criteria-petroleum-substances.

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.

Kimber, I., R.J. Dearman, D.A. Basketter, C.A. Ryan and G.F. Gerberick (2002), The local lymph node assay: past, present and future, *Contact Dermatitis*, Vol. 47, No. 6, pp. 315-328.

Klimisch, H. J., M. Andreae and S. Tillmann (1997), A Systematic Approach for Evaluating the Quality of Experimental Toxicological and Ecotoxicological Data, *Regulatory Toxicology And Pharmacology*, Vol. 25, pp. 1-5.

Mekenyan, O., S. Dimitrov, R. Serafimova, E. Thompson, S. Kotov, N. Dimitrova and J.D. Walker (2004), Identification of the structural requirements for mutagenicity by incorporating molecular flexibility and metabolic activation of chemicals I: TA100 model, *Chemical Research in Toxicology*, Vol. 17, No. 6, pp. 753-766.

Mekenyan, O., M. Todorov, R. Serafimova, S. Stoeva, A. Aptula, R. Finking and E. Jacob (2007), Identifying the structural requirements for chromosomal aberration by incorporating molecular flexibility and metabolic activation of chemicals, *Chemical Research in Toxicology*, Vol. 20, No. 12, pp. 1927-1941.

Mekenyan, O., G. Patlewicz, G. Dimitrova, C. Kuseva, M. Todorov, S. Stoeva, S. Kotov and E. Donner (2010), Use of genotoxicity information in the development of integrated testing strategies (ITS) for skin sensitization, *Chemical Research in Toxicology*, Vol. 18, No. 10, pp. 1519-1540.

Mekenyan, O. G., P. Petkov, S. Kotov, S. Stoeva, V. B. Kamenska, S. Dimitrov, M. Honma, M. Hayashi, R. Benigni, E. M. Donner and G. Y. Patlewicz (2012), Investigating the relationship between in vitroin vitro- in vivoin vivo genotoxicity: Derivation of mechanistic QSAR models for in vivo liver genotoxicity and in vivo bone marrow micronucleus formation which encompass metabolism, *Chemical Research in Toxicology*, Vol. 25, No. 2, pp. 277-296.

MERAG (2007), Metals Risk Assessment Guidance fact sheets.

Merrick, B.A. and M.E. Bruno (2004), Genomic and proteomic profiling for biomarkers and signature profiles of toxicity, *Current Opinion in Molecular Therapeutics*, Vol. 6, pp. 600-607.

Mitchell, P.C.H. (2009), Speciation of Mo compounds in water. Ultraviolet spectra and REACH read-across report. . Unpublished report for the International Molybdenum Association

NAFTA (2012), Technical working group on pesticides (TWG) (Quantitative) structure activity relationship

[(Q)SAR] guidance Document
http://www.epa.gov/oppfead1/international/naftatwg/guidance/qsar-guidance.pdf

National Library of Medicine website; http://chem.sis.nlm.nih.gov/chemidplus

Netzeva, T.I., M. Pavan and A.P. Worth. (2008), Review of (quantitative) structure-activity relationships for acute aquatic toxicity. *QSAR Combinatorial Science*, Vol. 27, pp. 77-90.

Niyogi, S. and C.M. Wood (2004), Biotic Ligand Model, a flexible tool for developing site-specific water quality guidelines for metals, *Environmental Science and Technology,* Vol. 38, pp. 6177-6192.

OECD Cooperative Chemicals Programme.
http://www.oecd.org/env/ehs/risk-assessment/oecdcooperativechemicalsassessmentprogramme.htm

OECD Existing Chemicals Database. www.oecd.org/env/existingchemicals/data

OECD Quantitative Structure-Activity Relationships Project:
(http://www.oecd.org/env/ehs/oecdquantitativestructure-activityrelationshipsprojectqsars.htm)

OECD (2000a), SIDS Initial Assessment Profile. Alkanes, C10-13, chloro-, No. 166, Series on Testing and Assessment. ENV/JM/MONO(2012)4/PART2, OECD, Paris.

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

OECD (2001a), SIDS Initial Assessment Report. Alpha Olefin Category, No. 166, Series on Testing and Assessment, ENV/JM/MONO(2012)4/PART3, OECD, Paris.

OECD (2001b), 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 (2003), SIDS Initial Assessment Profile. Xylenes Category No. 166, Series on Testing and Assessment. ENV/JM/MONO(2012)4/PART4, OECD, Paris.

OECD (2004a), OECD Principles for the Validation, for Regulatory Purposes, of (Quantitative) Structure-Activity Relationship Models.

OECD (2004b), SIDS Initial Assessment Profile. Ethylene glycols Category. No. 166, Series on Testing and Assessment, ENV/JM/MONO(2012)4/PART4, OECD, Paris.

OECD (2004c), SIDS Initial Assessment Profile, Butenes. 19-22 October 2004.
http://webnet.oecd.org/Hpv/ui/handler.axd?id=bea7ff32-f169-449e-8dcd-3fada777b36b

OECD (2004d), SIDS Initial Assessment Profile. High molecular weight phthalate esters. No. 166, Series on Testing and Assessment. ENV/JM/MONO(2012)4/PART4, OECD, Paris.

OECD (2005), SIDS Initial assessment profile zinc metal, zinc oxide, zinc distearate, zinc chloride, zinc sulphate, trizinc bis orthophosphate, No. 166, Series on Testing and Assessment. ENV/JM/MONO(2012)4/PART5, OECD, Paris.

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), SIDS Initial Assessment Profile. Long chain alcohols (C6-22 primary aliphatic alcohols, No. 166, Series on Testing and Assessment. ENV/JM/MONO(2012)4/PART5, OECD, Paris.

OECD (2006c), SIDS Initial Assessment Profile. Alkylamidopropyl betaines, No. 166, Series on Testing and Assessment. ENV/JM/MONO(2012)4/PART5, OECD, Paris.

OECD (2006d), SIDS Initial Assessment Profile Amine oxides, No. 166, Series on Testing and Assessment. ENV/JM/MONO(2012)4/PART5, OECD, Paris.

OECD (2006e), OECD Harmonised Templates, http://www.oecd.org/ehs/templates/#d.en.192217

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), SIDS Initial Assessment Profile. Alkyl sulfates, alkane sulfonates and α-olefin sulfonates, No. 166, Series on Testing and Assessment. ENV/JM/MONO(2012)4/PART5, OECD, Paris .

OECD (2007c), SIDS Initial Assessment Profile. Nitrates, No. 166, Series on Testing and Assessment. ENV/JM/MONO(2012)4/PART5, OECD, Paris.

OECD (2009a), The Guidance Document for Using the OECD (Q)SAR Application Toolbox to Develop Chemical Categories According to the OECD Guidance on Grouping Chemicals, No. 102. Series on Testing and Assessment, ENV/JM/MONO(2009)5, OECD, Paris.

OECD (2009b), Report of the Workshop on Structural Alerts for the OECD (Q)SAR Application Toolbox, No.101. Series on Testing and Assessment, ENV/JM/MONO(2009)4, OECD, Paris.

OECD (2009c), SIDS Initial Targeted Assessment Profile (Environment), Anthracene oils, No. 166, Series on Testing and Assessment. ENV/JM/MONO(2012)4/PART6, OECD, Paris.

OECD (2010a), SIDS Initial Assessment Profile. C2-C4 Aliphatic thiols category, No. 166, Series on Testing and Assessment. ENV/JM/MONO(2012)4/PART6, OECD, Paris.

OECD (2010b), SIDS Initial Assessment Profile. Alkyl chlorosilanes (Chlorotrimethylsilane, Dichlorodimethylsilane, Trichloromethylsilane, and Trichloroethylsilane) category, No. 166, Series on Testing and Assessment. ENV/JM/MONO(2012)4/PART6, OECD, Paris.

OECD (2010c), SIDS Initial Assessment Profile. Chloroformates, No. 166, Series on Testing and Assessment. ENV/JM/MONO(2012)4/PART6, OECD, Paris.

OECD (2011a), Report of the workshop on using mechanistic information in forming chemical categories, No. 138, Series on Testing and Assessment, ENV/JM/MONO(2011)8, OECD, Paris..

OECD (2011b), Manual for the Assessment of Chemicals
http://www.oecd.org/env/ehs/risk-assessment/manualfortheassessmentofchemicals.htm

OECD (2011c), SIDS Initial Assessment Profile. C1-13 Primary Amines, No. 166, Series on Testing and Assessment. ENV/JM/MONO(2012)4/PART6, OECD, Paris.

OECD (2011d), SIDS Initial Assessment Profile. C14-C20 Aliphatic hydrocarbon solvents, No. 166, Series on Testing and Assessment. ENV/JM/MONO(2012)4/PART6, OECD, Paris.

OECD (2012a), The Adverse Outcome Pathway for Skin Sensitisation Initiated by Covalent Binding to Proteins, Series on Testing and Assessment, No. 168, OECD, Paris.

OECD (2012b), SIDS Initial Assessment Profile. C10-C13 Aromatic Hydrocarbon Solvents Category.
http://webnet.oecd.org/Hpv/ui/handler.axd?id=884a80e0-a657-48cd-8748-cd361a682d06

OECD (2012c), SIDS Initial Assessment Profile. Tertiary Amines,

http://webnet.oecd.org/HPV/UI/handler.axd?id=ee5efcff-697d-48f7-ab7d-032356a83370

OECD (2012d) Task Force on Hazard Assessment. Progress with the development of guidance on the identification of substances of unknown and variable composition or of biological origin (UVCBs), Report ENV/JM/HA(2012)6. Organisation for Economic Co-operation and Development, Paris, France.

OECD (2012e), Report of an OECD workshop on metals specificities in environmental risk assessment 7-8 September 2011, Paris, France Series on Testing & Assessment No. 165. ENV/JM/MONO(2012)2.

OECD (2012f), OECD Environment, Health and Safety Publications Series on the Safety of Manufactured Nanomaterials, No. 33. Important issues On Risk Assessment Of Manufactured Nanomaterials. Report ENV/JM/MONO(2012)8, Inter-Organization Programme for the Sound Management of Chemicals (IOMC), Organisation for Economic Co-operation and Development (OECD), Paris, France.

OECD (2012g), OECD Environment, Health and Safety Publications Series on the Safety of Manufactured Nanomaterials, No. 36. Guidance on Sample Preparation and Dosimetry for the Safety Testing of Manufactured Nanomaterials. Report ENV/JM/MONO(2012)40, Inter-Organization Programme for the Sound Management of Chemicals (IOMC), Organisation for Economic Co-operation and Development (OECD), Paris, France.

OECD (2013a), Guidance Documents and Training Materials for Using the Toolbox. http://www.oecd.org/env/ehs/risk-assessment/theoecdqsartoolbox.htm#Guidance_Documents_and_Training_Materials_for_Using_the_Toolbox

OECD (2013b), Guidance document on developing and assessing adverse outcome pathways, No. 184, Series on Testing and Assessment. ENV/JM/MONO(2013)6, OECD, Paris.

OECD (2014a), OECD guidance for characterising oleochemical substances (in prep)

OECD (2014b), Weight of evidence assessment for skin sensitisation potential for 4-isopropylaniline (in prep)

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

Patlewicz, G,Y., W. El-Deredy (1999), A QSAR - Bayesian Neural network model to identify molecular properties causing eye irritation in cationic surfactants. In G.C. Gini and A.R. Katritzky (editors), Predictive Toxicology of Chemicals: Experiences and Impact of AI Tools - AAAI Spring Symposium, AAAI Press, Menlo Park, CA.

Patlewicz, G., D.A. Basketter, C.K. Smith, S.A.M. Hotchkiss and D.W. Roberts (2001), Skin-sensitization structure-activity relationships for aldehydes, *Contact Dermatitis*, Vol. 44, No. 6, pp. 331-336.

Patlewicz, G., S. Dimitrov, L.K. Low, P.S. Kern, G.D. Dimitrova, M.I. Comber, A.O. Aptula, R.D. Phillips, J. Niemelä, C. Madsen, E.B. Wedebye, D.W. Roberts, P.T. Bailey and O.G. Mekenyan (2007), TIMES-SS - A promising tool for the assessment of skin sensitization hazard. A characterization with respect to the OECD validation principles for (Q)SARs and an external evaluation for predictivity, *Regulatory Toxicology and Pharmacology*, Vol. 48, No. 2, pp. 225-239.

Patlewicz G, N. Ball, E.D. Booth, E. Hulzebos, E. Zvinavashee, and C. Hennes (2013), Use of Category approaches, Read-across and (Q)SAR: General considerations, *Regulatory Toxicology and Pharmacology,* in press

PETROTOX (2011), PETROTOX: A tool for the hazard assessment of petroleum substances, Version 3.06. Conservation of Clean Air and Water in Europe (CONCAWE). http://www.concawe.be/content/default.asp?PageID=778 .

Regulation (EC) No 1272/2008 of the European Parliament and of the Council of 16 December 2008 on classification, labelling and packaging of substances and mixtures (http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CONSLEG:2008R1272:20110419:EN:PDF) amending and repealing the Dangerous Substances Directive 67/548/EE and the Dangerous Preparations Directive 1999/45/EC, and amending Regulation (EC) No 1907/2006.

Resins, U.E.A. (2003), Guide to the classification and labelling of UV/EB Acrylates. CEFIC, Brussel

Russom, C.L., S.P. Bradbury, S.J. Broderius, D.E. Hammermeister and R.A. Drummond (1997), Predicting Modes of Toxic Action from Chemical Structure: Acute Toxicity in the Fathead Minnow (*Pimephales promelas*), *Environmental Toxicology and Chemistry*, Vol. 16, No. 5, pp. 948-967.

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

Saghir, S.A., M.J. Bartels, D.L. Rick, A.T. McCoy, R.J. Rasoulpour, R.G. Ellis-Hutchings, M. Sue Marty, C. Terry, J.P. Bailey, R. Billington and J.S. Bus (2012), Assessment of diurnal systemic dose of agrochemicals in regulatory toxicity testing--an integrated approach without additional animal use, *Regulatory Toxicology and Pharmacology*, Vol. 63, No. 2, pp. 321-332.

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.

SciFinder http://www.cas.org/SCIFINDER

Schoeters, I. and V. Verougstraete (2007), Case studies on inorganic metal compounds, 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. .

Schultz, T.W. (2010), Adverse Outcome Pathways: a way of linking chemical structure to in vivo toxicological hazards. In silico Toxicology, Principles and Applications, Cronin& Madden. The Royal Society of Chemistry, Chapter 14, (7).

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

Schultz, T.W., T.I. Netzeva, D.W. Roberts and M.T.D. Cronin (2005), Structure-Toxicity Relationships for the Effects to Tetrahymena pyriformis of Aliphatic, Carbonyl-Containing, α,β-Unsaturated Chemicals, *Chemical Research in Toxicology*, Vol. 18, No. 2, pp. 330–341.

Serafimova, R., M. Todorov, T. Pavlov, S. Kotov, E. Jacob, A. Aptula and O. Mekenyan (2007), Identification of the structural requirements for mutagencitiy, by incorporating molecular flexibility and metabolic activation of chemicals. II, *General Ames mutagenicity model,* Vol. 20, No. 4, pp. 662-676.

Smith S.W. (2009) REVIEW: Chiral toxicology: It's the same thing…Only different, *Toxicological Sciences*, Vol. 101, No. 1, pp. 4-30

Trimmer, G.W. (2008), Whole-Body Inhalation Developmental Toxicity Study in Rats with Baseline Gasoline Vapor Condensate (MRD-00-695), ExxonMobil Biomedical Sciences, Inc. Study No

169534; sponsored by the American Petroleum Institute

Tsakovska, I. and A. Worth (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 inorganic metal REACH Guidance on Chemical Categories and Read Across, Worth A. & Patlewicz G. (Eds.), EUR report no 22481 EN.

UNEP Chemicals, (2006), http://www.chem.unep.ch/

United Nations (2013), Globally Harmonised System of Classification and Labelling of Chemicals. http://www.unece.org/trans/danger/publi/ghs/ghs_rev05/05files_e.html

US EPA Identification Methodology http://www.epa.gov/oppt/sf/tools/aim.htm

Van der Weiden, M.E.J., J. van der Kolk, R. Bleumink, W. Seinen, and M. van den Berg (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*, Vol. 24, pp.123- 142.

Verhaar, H.J.M., C.J. van Leeuwen and J.L.M. Hermen (1992), Classifying environmental pollutants. 1. Structure-activity relationships for prediction of aquatic toxicity, *Chemosphere*, Vol. 25, pp. 471-491.

Walker, M.K., and R.E. Peterson (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,* Vol. 21, pp. 219-238.

Worth, A. and G. Tier (2007), A Compendium of Case Studies that helped to shape the REACH Guidance on Chemical Categories and Read Across, EUR report no 22481 EN. Worth A & Patlewicz G (Eds).

Wu, S., K. Blackburn, J. Amburgey, J. Jaworska and T. Federle (2010), A framework for using structural, reactivity, metabolic and physicochemical similarity to evaluate the suitability of analogs for SAR-based toxicological assessments, *Regulatory Toxicology and Pharmacology,* Vol. 56, pp. 67-81.

Zabel, E. W., P.M. Cook and R.E. Peterson (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***,** Vol 31, pp. 315-328.

9 LIST OF ABBREVIATIONS

AD	Applicability Domain
AOP	Adverse outcome pathway
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	Conservation of Clean Air and Water in Europe: 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
KE	Key event
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
MIE	Molecular initiating event
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
PBPD	Physiologically based Pharmacodynamic
PBPK	Physiologically based Pharmacokinetic
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
Substance	
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

10 APPENDIX

Table 14. Specific aspects of endpoint read-across justifications[58]

Endpoint	Approaches and tools
Physical-chemical parameters	Physical-chemical parameters play a critical role in addressing many aspects of the substance's behaviour and in characterising the chemical similarity for read-across purposes. Basic physical-chemical properties provide key information for the assessment of a chemical and in particular for the assessment of the environmental properties. Consequently experimental data or valid QSAR predictions should normally be available (or should be reasonably obtainable). However, there may occasionally be practical problems, especially for UVCBs, when the use of read-across techniques will be required. Vapour pressure, logKow, water solubility, molecular weight, pKa are critical when considering bioaccumulation in the environment and absorption in the animal/human organism and should be addressed for the category members. Tools: EPI Suite, ACD/ Percepta, OECD QSAR Toolbox, ChemAxon, T.E.S.T.
Aquatic toxicity	To facilitate and justify the read-across approach for the aquatic endpoint toxicity, tools like rule base schemes, (Q)SARS-based and WoE approaches are helpful to indicate the mode of action of the substance, and elements that can be used to demonstrate similarity between two or more analogues. A combination of (Q)SAR model and measured data might also be considered to test any hypothesis and strengthen the overall strategy. Tools: ECOSAR, Aquatic toxicity classification by ECOSAR, TOPKAT, T.E.S.T., TIMES , Verhaar rulebase within Toxtree or as a profiler within the OECD (Q)SAR Toolbox, OASIS Acute Aquatic Toxicity Mode of Action Profiler within the OECD QSAR Toolbox, LAZAR Danish QSAR database with predictions from DTU developed models for acute toxicity to Fathead minnow, Daphnia magna, Pseudokirchneriella subcapitata and Tetrahymena pyriformis.
Biodegradation	Biodegradation is a critical endpoint as it impacts upon classification and labelling PBT assessment

[58] ECETOC TR116 is acknowledged as a significant source of information for this table

Endpoint	Approaches and tools
	Waivers for hydrolysis and adsorption-desorption testing. Building a WoE case for bioaccumulation. Consequently any read-across strategy needs to be robust. Several databases offer a great number of biodegradation pathways, but if other data is available, this source of information should only be used as a part of a weight-of-evidence approach. For experimental studies, the protocol that has been used needs to be evaluated to ensure that it is fit for purpose for the particular substance and investigation. (Q)SAR modelling may be useful, but it must be interpreted appropriately to ensure that it is correlated to the physical-chemical properties of the substance. This approach is applied mostly to substances that are not readily biodegradable substances rather than biodegradable substances (ECHA website 2012). Tools: Biowin, TOPKAT, Catalogic, Danish QSAR predictions database.
Bioaccumulation	The bioconcentration factor (BCF), bio-magnification factor (BMF) and bioaccumulation factor (BAF) are used in bioaccumulation assessments- which are quantitative. Read-across can be applied if a substance has a valid BCF for a structurally close related substance. When (Q)SAR models and common databases are used to provide BCF values, properties like ionisation, hydrolysis, adsorption, molecular mass and size data, and degradation need to be considered. Available experimental studies should be used in a read-across approach. In this case, it is important to select results from studies with a relevant protocol. A weight of evidence strategy can be also used to strengthen and support the read-across approach for a category using all available information that can contribute the potential for bioaccumulation (data from model, ADME, *in vitro* and *in vivo* assays). Tools: Catalogic, BCFBAF, Caesar (VEGA) Model for BCF, T.E.S.T.
Mammalian Toxicity	
Acute Oral Route	There are very many modes of action for acute oral toxicity making modelling difficult. The evaluation of structural similarity in terms of functional group, physical-chemical profile, and steric and hydrophobic moieties are key elements in developing a read-across approach. Metabolism data are very useful to demonstrate common metabolites, but if not already available are likely to be impractical due to the relative cost/benefit of the data. In case of a read-across based or category based on strong structural similarity. Read-across approach can be based upon chemical state, such as hydrolysis or ionisation of salts with data to demonstrate the likely bioavailability of a substance, and its common nature, under physiological conditions. The read-across approach can be supported by results from *in vitro* cytotoxicity assays, such as the neutral red uptake assay. Tools: T.E.S.T., TOPKAT ,OECD QSAR Toolbox
Acute	The physical-chemical properties of the substance and chemical reactivity are major determinants of toxicity. Particle size, vapour pressure, and water solubility are all

Endpoint	Approaches and tools
Inhalation route	important especially in the case of volatile substances, both solid aerosol and liquid aerosols may be respirable and trigger hazard. A read-across approach of volatile substances could consider information from other endpoints where narcosis and electrophilic reactivity play a role. In this case, (QSAR) models can be used to demonstrate whether a general narcosis mode of action was relevant. For non-volatiles substances, the read-across approach can be supported by information from other routes of exposure, i.e., dermal and oral and there are default models available. Tools: TOPKAT, OECD QSAR Toolbox.
Irritation Skin	Physical-chemical information is important for evaluation of the endpoint - especially information on pH, where low and high values are sufficient to determine a substances likely skin and eye irritant / corrosive potential, e.g., pH <2 or >11.5 are considered as corrosive. In order to predict the absence of skin corrosion/irritation, the read-across approach can be supported using SARs models, together with the physical-chemical properties of the substance. Danish QSAR database with predictions from DK DTU developed model for severe vs. mild skin irritation.
Eye	The same considerations as for skin irritancy and physical-chemical parameters apply. Eye corrosion/irritation can be demonstrated by using SARs models and the physical-chemical properties of the substance. Alternative *In vitro* methods are available to support read-across. Tools: TOPKAT, BfR rulebase within Toxtree, Eye and Skin irritation inclusion and exclusion rules by BfR Profiler with the OECD QSAR Toolbox, Derek Nexus.
Skin sensitisation	A weight of evidence strategy can be applied using *in vivo* and *in vitro* data, and evaluating the physical-chemical profile of the substance. EURL ECVAM has recently published its Strategy for Replacement of Animal Testing for Skin Sensitisation Hazard Identification and Classification. Databases, QSAR models may be used to identify additional substances to use in the read-across approach based on existing data and similarity in physical-chemical structure. (Q)SAR models and *in vitro* methods are focused mainly on reactivity = Protein binding, metabolic activation, internalisation and processing by Langerhans cells (LC), transport of antigen by LC to lymph node and activation of T lymphocyte. Tools: CAESAR Model for Skin Sensitisation, TIMES, TOPKAT, Derek Nexus, MCASE, SMARTS alerts within Toxtree, Protein binding profilers within the OECD QSAR Toolbox: Protein binding by OECD profiler, Protein binding by OASIS v1.1 profiler and Protein binding alerts for skin sensitisation by OASIS v1.1 .

Endpoint	Approaches and tools
	Danish QSAR database contains predictions from the commercial MultiCASE model A33.
Genotoxicity	Data from *in vitro* tests are usually available or should be acquired for strategic category members. A large body of experimental data on mutagenicity has allowed identification of structural alerts for mutagenicity, and can be considered within a category as part of the supporting evidence. (Q)SAR models are able to make prediction of mutagenicity testing (Salmonella) and in vivo mutagenicity and may provide supporting evidence. Tools: CAESAR, TIMES, TOPKAT, Derek Nexus, MCASE, T.E.S.T., LAZAR, Leadscope Model Applier, Benigni/Bossa Rulebase within Toxtree, ToxMIC-ISS plug-in allows the identification of Structure Alerts for the in vivo micronucleus assay within Toxtree Profilers within the OECD QSAR Toolbox: DNA binding by OASIS v1.1 profiler and DNA Alerts for Ames, CA and MN by OASIS, DNA Binding by OECD profiler, In vitro, in vivo and Carcinogenicity/mutagenicity alerts by ISS. Danish QSAR database (a range of genotox endpoints both in vivo and in vitro, are covered, also carcinogenicity in rat/mouse in male/female covered); the database contains predictions from commercial and DTU developed models.
Repeated Dose Toxicity	The source chemical study(ies) need to be reviewed for fit for purpose for the read-across considering, test material, route of exposure, test species, study type and validity, protocol, extent of observations, in order to determine if a study is a suitable source for read-across within a category. The similarity in the toxicological profile across all the human health endpoints will be considered as well as ADME information on informing on a suitability of the read-across to target chemicals. At the moment, there are no *in vitro* methods available that can replace *in vivo* repeated dose toxicity data. However, the SEURAT-1 Research Initiative of the EU funded within the 7th Framework Programme is currently conducting six complementary research projects aimed at ultimately replacing animals in repeated dose toxicity testing. (Q)SAR models might be used to demonstrate similarity in reactivity and support existing data. Tools: TOPKAT (LOAEL, MDT,) Derek Nexus, LAZAR, HESS Profiler within the OECD QSAR Toolbox, NEDO/METI Project, Fraunhofer database.
Reproductive and developmental Toxicity	In building a hypothesis on reproductive toxicity, structural, functional as well as ADME considerations will be used. Gross pathology and histopathology data from repeated dose toxicity studies should also be used. As reproductive and developmental toxicity are complex endpoints and the mode of action is not usually known, currently read-across must be based upon experimental data. Screening studies (OECD 421 and 422) on category members can provide useful data and provide confidence in the read-across of higher tier studies from source chemicals and test the validity any read-across hypothesis for the endpoint. (Q)SAR models have been developed and may be useful for supporting trends seen in existing data for the category members.

Endpoint	Approaches and tools
	Currently *in vitro* assays, only can be used support specific outcomes on reproductive organs. However, a "Feasibility Study" which concluded the ReProTect an Integrated Project of the EU (funded within the 6th Framework Programme) identified a test battery of 14 *in vitro* assays that allowed a robust prediction of adverse effects on fertility and embryonic development of 10 test chemicals with toxicologically well-documented profiles *in vivo*. Tools: Derek Nexus, TOPKAT, CASEAR Model for Developmental Toxicity, TIMES, Leadscope Model Applier, MCASE, rtER expert system developed by EPA as well as the associated ER binding profiler as encoded within the OECD QSAR Toolbox. Danish QSAR database contains predictions for EST receptor binding and activation, AR receptor activation as well as teratogenicity (predictions from a commercial MultiCASE model based on human training set data).

Table 15. References and context of use for prediction tools

Endpoint	Name of Tool	Reference	Context of Use
Physical-chemical parameters	EPI Suite	US EPA http://www.epa.gov/opptintr/exposure/pubs/episuite.htm	Estimates parameters such as logKow, melting point, boiling point, vapour pressure, water solubility
	ACD/Percepta	ACD/Labs http://www.acdlabs.com/products/percepta/	Estimator of parameters including water solubility, boiling point, logKow, LogD, pKa
	OECD (Q)SAR Toolbox	OECD http://www.oecd.org/chemicalsafety/risk-assessment/theoecdqsartoolbox.htm	Encodes the EPISuite predictors. Also contains available experimental data on key physchem parameters
	ChemAxon	ChemAxon http://www.chemaxon.com/	Predictors for logKow, LogD, pKa
	T.E.S.T.	US EPA http://www.epa.gov/nrmrl/std/qsar/qsar.html#TEST	Physical property endpoints include boiling point, flash point, surface tension, viscosity, density, water solubility, and thermal conductivity
Aquatic toxicity	ECOSAR	US EPA http://www.epa.gov/opptintr/exposure/pubs/episuite.htm	Main utility is to predict acute and chronic effects for fish, daphnia and algae. Structured in the form of SARs for specific chemical classes which provide some indication of likely MOA.

Endpoint	Name of Tool	Reference	Context of Use
	Aquatic toxicity classification by ECOSAR Profiler within the OECD QSAR Toolbox	OECD http://www.oecd.org/chemicalsafety/risk-assessment/theoecdqsartoolbox.htm	Profiler to help assign MOA on the basis of chemical class for the purposes of deriving endpoint specific chemical categories
	TOPKAT	Accelrys http://accelrys.com/	Global models to predict acute toxicity to fish and daphnia.
	T.E.S.T.	US EPA http://www.epa.gov/nrmrl/std/qsar/qsar.html#TEST	Models to predict 96-hr fathead minnow LC50, 48-hr Daphnia magna LC50, Tetrahymena pyriformis 50% IGC50
	TIMES	LMC http://oasis-lmc.org/products/software/times.aspx	Available models include those to predict 96-hr fathead minnow LC50, 48-hr Daphnia magna LC50, Tetrahymena pyriformis 50% IGC50
	OECD (Q)SAR Toolbox	OECD http://www.oecd.org/chemicalsafety/risk-assessment/theoecdqsartoolbox.htm	Contains available experimental data to help develop new trend analysis for the prediction of key aquatic endpoints
	Verhaar rulebase; Acute aquatic classification by Verhaar profiler	JRC and OECD http://www.oecd.org/chemicalsafety/risk-assessment/theoecdqsartoolbox.htm	The Verhaar rulebase enables substances to be characterised according to their likely MOA. The scheme has been implemented into Toxtree as well as the OECD QSAR Toolbox
	OASIS Acute Aquatic Toxicity Mode of Action Profiler	OECD http://www.oecd.org/chemicalsafety/risk-assessment/theoecdqsartoolbox.htm	A scheme to enable substances to be categorised according to their likely MOA.
	Lazy Structure Activity Relationships (LAZAR)	http://lazar.in-silico.de/predict	Model to predict 96 hr fathead minnow LC50
	Danish QSAR predictions database	http://qsar.food.dtu.dk	Predictions made by DTU models in MultiCASE MC4PC software for short-term toxicity to Fathead minnow (96h LC50), Daphnia magna (48h LC50), Pseudokirchneriella subcapitata (EC50) and Tetrahymena pyriformis (46h

Endpoint	Name of Tool	Reference	Context of Use
			IGC50).
Biodegradation	Biowin	US EPA http://www.epa.gov/opptintr/exposure/pubs/episuite.htm	Models to predict ready biodegradation
	TOPKAT	Accelrys	Global model to predict ready biodegradability
	Catalogic	http://oasis-lmc.org/products/software/catalogic.aspx	Contains a suite of models to predict biodegradation under different study protocols e.g. OECD 301C While accounting for microbial metabolism
	Danish QSAR predictions database	http://qsar.food.dtu.dk	Predictions made by DTU model in MultiCASE MC4PC software for 165.000 chemicals for MITI ready / not ready.
Bioaccumulation	Catalogic	LMC http://oasis-lmc.org/products/software/catalogic.aspx	Contains BCF base line model which predicts BCF While accounting for modulating factors such as metabolism, size, ionisation
	BCFBAF	US EPA http://www.epa.gov/opptintr/exposure/pubs/episuite.htm	Models to predict BCF and BAF
	CAESAR Model for BCF	CAESAR http://www.vega-qsar.eu/index.php	Global model for BCF, now part of the VEGA platform
	T.E.S.T.	US EPA http://www.epa.gov/nrmrl/std/qsar/qsar.html#TEST	Software encodes the CAESAR BCF model
Mammalian toxicity: Acute toxicity	T.E.S.T.	US EPA http://www.epa.gov/nrmrl/std/qsar/qsar.html#TEST	Global model for the prediction of rat LD50
	TOPKAT	Accelrys http://accelrys.com/	Global model for the prediction of rat LD50
	OECD (Q)SAR Toolbox	OECD http://www.oecd.org/che	Contains experimental data of LD50 in rodents

Endpoint	Name of Tool	Reference	Context of Use
		micalsafety/risk-assessment/theoecdqsartoolbox.htm	
Mammalian toxicity: eye	TOPKAT	Accelrys http://accelrys.com/	Global model to discriminate eye irritation potency
	BFR Rulebase within Toxtree	JRC	Scheme to classify eye irritants on the basis of structural alerts and to assign no classification for substances that meet specific physicalchemical parameter thresholds
	Eye irritation inclusion and exclusion rules by BfR Profiler within the OECD QSAR Toolbox	OECD http://www.oecd.org/chemicalsafety/risk-assessment/theoecdqsartoolbox.htm	Permits a categorisation of substances based on presence of alerts and extremes of physicalchemical parameters
	Derek Nexus	Lhasa Limited http://www.lhasalimited.org/products/derek-nexus.htm	Contains SAR for eye irritation that is useful to characterise MOA information
Mammalian toxicity: Skin irritation	Danish QSAR predictions database	http://qsar.food.dtu.dk	Predictions made by DTU model in MultiCASE MC4PC software for 165.000 chemicals for severe vs. mild skin irritation.
Mammalian toxicity: skin sensitization	CAESAR Model for Skin Sensitisation	CAESAR http://www.vega-qsar.eu/index.php	Global model for sensitisation, now part of the VEGA platform
	TIMES	Patlewicz *et al*., 2007	Hybrid expert system to predict sensitisation potency While accounting for metabolism
	TOPKAT	Accelrys http://accelrys.com/	Global model for the prediction of sensitising potency
	Derek Nexus	Lhasa Limited http://www.lhasalimited.org/products/derek-nexus.htm	Contains SARs for sensitisation which are helpful to assign MOA
	MCASE	Multicase Inc. http://www.multicase.com/products/prod01.htm	Global model for prediction of sensitisation

Endpoint	Name of Tool	Reference	Context of Use
	Protein binding by OECD profiler within the OECD QSAR Toolbox	OECD http://www.oecd.org/chemicalsafety/risk-assessment/theoecdqsartoolbox.htm	Alerts which characterise electrophilic reactivity based on organic chemistry principles. Assigns substances into reaction mechanistic domains that are pertinent for the assessment of skin sensitisation potential
	Protein binding by OASIS v1.1 profiler within the OECD QSAR Toolbox Protein binding alerts for skin sensitisation by OASIS v1.1 within the OECD QSAR Toolbox	OECD http://www.oecd.org/chemicalsafety/risk-assessment/theoecdqsartoolbox.htm	Mirror the SARs contained within the TIMES skin sensitisation model
	SMARTS alerts within Toxtree	JRC	Assignment of reaction mechanistic domains. Based on the reaction principles defined by Aptula and Roberts,(2006)
	Danish QSAR predictions database	http://qsar.food.dtu.dk	Predictions made by MultiCASE MC4PC (commercial A33) model for 165.000 chemicals for allergic contact dermatitis.
Mammalian toxicity: genotoxicity	CAESAR	CAESAR http://www.vega-qsar.eu/index.php	Global model for Ames mutagenicity, now part of the VEGA platform
	TIMES	Mekenyan *et al*., 2012; Mekenyan *et al*., 2004; Mekenyan *et al*., 2007; Serafimova *et al*., 2007	Hybrid expert system which contains a suite of models for the prediction of Ames mutagenicity, *in vitro* chromosomal aberration, *in vivo* liver genotoxicity and *in vivo* micronucleus. All models account for metabolism
	TOPKAT	Accelrys http://accelrys.com/	Global model for Ames mutagenicity
	Derek Nexus	Lhasa Limited http://www.lhasalimited.org/products/derek-nexus.htm	SAR for mutagenicity, chromosomal aberration, DNA damage etc Useful to categorise chemicals on the basis of their likely MOA
	MCASE	Multicase Inc. http://www.multicase.co	Global model for Ames mutagenicity

Endpoint	Name of Tool	Reference	Context of Use
		m/products/prod01.htm	
	Leadscope Model Applier	Leadscope https://www.leadscope.com/model_appliers/	Global models for a range of different genetic toxicity endpoints
	DNA binding by OASIS v1.1 profiler within the OECD QSAR Toolbox DNA Alerts for Ames, CA and MN by OASIS within the OECD QSAR Toolbox	OECD http://www.oecd.org/chemicalsafety/risk-assessment/theoecdqsartoolbox.htm	Mirror the SARs contained within the TIMES
	DNA Binding by OECD profiler within the OECD QSAR Toolbox	http://www.oecd.org/chemicalsafety/risk-assessment/theoecdqsartoolbox.htm	Alerts which characterise electrophilic reactivity based on organic chemistry principles
	Benigni/Bossa Rulebase within Toxtree	JRC	A compilation of SARs made by R Benigni and C Bossa
	ToxMIC-ISS plug-in allows the identification of Structure Alerts for the *in vivo* micronucleus assay within Toxtree	JRC	A compilation of SARs for *in vivo* MN
	In vitro, *in vivo* and Carcinogenicity/mutagenicity alerts by ISS within the OECD QSAR Toolbox	OECD http://www.oecd.org/chemicalsafety/risk-assessment/theoecdqsartoolbox.htm	An update and refinement of the Benigni-Bossa and Toxx-MIC rulebases re-coded within the Toolbox
	T.E.S.T.	US EPA http://www.epa.gov/nrmrl/std/qsar/qsar.html#TEST	Global model for Ames mutagenicity
	Lazy Structure Activity Relationships (LAZAR)	http://lazar.in-silico.de/predic t	Global models based on the following datasets DSSTox Carcinogenic Potency DBS and Kazius-Bursi Salmonella

Endpoint	Name of Tool	Reference	Context of Use
	Danish QSAR predictions database	http://qsar.food.dtu.dk	Predictions made by MultiCASE MC4PC software for 165.000 chemicals for: Ashby fragments (commercial) In vitro mutagenicity • Ames test (commercial) • Ames sub-models (DTU models for S9 activation, Base pair mutation, Frame shift mutation, Potency at least 10 times over control group) • Chromosomal aberration CHO (commercial model) • Mouse lymphoma TK assay (DTU model) • Unscheduled DNA synthesis (DTU model) • CHO/HGPRT forward mutation assay (DTU model) • SHE cell transformation (DTU model) In vivo mutagenicity • Drosophila sex-linked recessive lethal (DTU model) • Mouse micronucleus bone marrow (DTU model) • Rodent dominant lethal (DTU model) • Mouse SCE bone marrow (DTU model) • Mouse Comet assay (DTU model)
Mammalian toxicity: repeated dose toxicity	TOPKAT (LOAEL, MTD)	Accelrys http://accelrys.com/	Global model for the prediction of LOAEL, MTD
	Derek Nexus	Lhasa Limited http://www.lhasalimited.org/products/derek-nexus.htm	SARs for many different endpoints associated with repeated dose toxicity
	OECD QSAR Toolbox	OECD http://www.oecd.org/chemicalsafety/risk-assessment/theoecdqsart	Contain repeated dose data

Endpoint	Name of Tool	Reference	Context of Use
		oolbox.htm	
	NEDO/METI Project	Hayashi, 2011	
	Fraunhofer	http://www.fraunhofer-repdose.de/	Database of repeated dose toxicity information. Also made available within the OECD QSAR Toolbox
	Lazy Structure Activity Relationships (LAZAR)	http://lazar.in-silico.de/predict	Global FDA v3b Maximum Recommended Daily Dose model
	Hazard Evaluation Support System Integrated Platform (HESS) Profiler	OECD http://www.oecd.org/chemicalsafety/risk-assessment/theoecdqsartoolbox.htm	Profiler within the OECD QSAR Toolbox to help assign MOA
	Hazard Evaluation Support System Integrated Platform (HESS)	http://www.safe.nite.go.jp/english/kasinn/qsar/hess-e.html	Expert system containing repeated dose toxicity information to facilitate hazard assessment through the development of chemical categories. System mimics the structure/platform of the OECD QSAR Toolbox
Mammalian toxicity: reproductive and developmental toxicity	Derek Nexus	Lhasa Limited http://www.lhasalimited.org/products/derek-nexus.htm	SARs for teratogenicity, developmental toxicity, reproductive effects
	TOPKAT	Accelrys http://accelrys.com/	Global model for developmental toxicity
	CAESAR Model for Developmental Toxicity	CAESAR http://www.vega-qsar.eu/index.php	Global model for developmental toxicity
	Leadscope Model Applier	Leadscope https://www.leadscope.com/model_appliers/	Sex specific global models for developmental toxicity and reproductive effects in rodents
	MCASE	Multicase Inc. http://www.multicase.com/products/prod01.htm	

Endpoint	Name of Tool	Reference	Context of Use
	rtER expert system developed by EPA as well as the associated ER binding profiler within the OECD QSAR Toolbox	OECD http://www.oecd.org/chemicalsafety/risk-assessment/theoecdqsartoolbox.htm	Encoded as a expert system within the Toolbox and as a profiler to assign chemicals on the basis of their likely ER MOA
	TIMES	http://oasis-lmc.org/products/software/times.aspx	Models for ER, and AR Binding affinities
Potential for metabolisation and potential metabolites	OECD QSAR TB (skin, liver, environmental,/ simulated or observed, / metabolites indicated but not probability/ freely downloadable from OECD web site) SMARTcyp: predicts the sites in molecules that are most liable to cytochrome P450 mediated metabolism Probability for reaction with CYPs Meta2print: As above & Identity and probability of metabolites generated, hased on both phase I and II reactions	http://www.oecd.org/chemicalsafety/risk-assessment/theoecdqsartoolbox.htm http://www.farma.ku.dk/smartcyp/ Copenhagen University http://www-metaprint2d.ch.cam.ac.uk/metaprint2d/about.html University of Cambridge/ Department of Chemistry & Unilever Centre for Molecular Science Informatics	Various in relation to when potential metabolisation is of significance for chemical categorization
	Danish QSAR predictions database	http://qsar.food.dtu.dk	Predictions made by MultiCASE MC4PC software for 165.000 chemicals for ER binding (DTU model), ER reporter gene (DTU model) and teratogenicity (commercial model).

Index

(bio)transformation product, 24
adverse effect, 31, 50, 51, 52
adverse outcome pathway, 12, 33, 49, 59
analogue approach, 9, 11, 13, 14, 15, 16, 19, 22, 23, 26, 27, 53, 58, 59, 61, 62, 63, 66, 79, 105, 106
AOP. *See* adverse outcome pathway
applicability domain, 16, 17, 22, 64, 69
bioavailability, 12, 20, 21, 23, 44, 47, 52, 53, 61, 76, 96, 97, 99, 100, 102, 103, 110, 111, 129
biotransformation, 20
boundaries. *See* boundary
boundary, 17, 35, 52, 53, 91
breakpoint, 18, 20, 52, 91
category approach, 12, 14, 15, 16, 18, 19, 23, 25, 26, 29, 32, 33, 36, 43, 45, 50, 53, 56, 60, 66, 67, 68, 74, 96, 98, 105, 112, 120
category definition, 16, 37, 68, 75, 92, 94, 105
category description, 16, 22, 91
category hypothesis, 23, 68, 70, 76
category justification, 9, 18, 28, 35, 36, 43, 69, 71, 74, 112, 113
cluster, 16
dross, 98
evidence, 11, 19, 23, 25, 26, 28, 29, 30, 38, 46, 48, 53, 54, 56, 57, 63, 71, 72, 73, 75, 93, 99, 100, 131
frits, 98
key event, 21, 50, 51, 72, 73
mechanism of action, 12, 23, 41, 49, 51, 61, 62, 68, 72, 74, 79, 97, 100, 104, 106, 111
metabolism, 12, 13, 18, 19, 20, 28, 33, 45, 46, 47, 48, 56, 57, 60, 67, 71, 76, 79, 80, 81, 82, 111, 134, 135, 136
mode of action, 26, 45, 49, 50, 51, 54, 57, 61, 63, 69, 97, 106, 111, 128, 129, 131
molecular initiating event, 17, 21, 22, 50, 72, 73
multi-constituent substances, 34
MCS, 35, 36, 82
QAAR, 24, 54, 55
QSAR
QSAR/SAR, 10, 13, 24, 25, 31, 45, 52, 55, 56, 57, 59, 60, 61, 63, 64, 66, 67, 70, 72, 74, 86, 91, 92, 93, 113, 128, 129, 130, 131, 133, 135, 136, 137, 139, 140
source chemical, 11, 15, 16, 19, 23, 24, 27, 29, 30, 31, 34, 35, 45, 50, 51, 52, 58, 59, 62, 63, 71, 105, 106, 131
subcategory, 16
subcategories, 18, 37, 39, 74
target chemical, 11, 13, 14, 16, 19, 23, 27, 29, 30, 34, 44, 45, 49, 55, 57, 58, 59, 61, 62, 63, 64, 71, 105, 106
trend analysis, 12, 18, 45, 52, 53, 55, 63, 69, 74, 86, 112, 133
UVCB, 12, 27, 30, 33, 34, 35, 36, 45, 57, 58, 82, 87, 91, 94, 95, 98, 110
weight of evidence, 13, 17, 23, 26, 28, 31, 35, 57, 99, 100, 109, 128, 130