RIVM report 601200009/2006

Workability of the guidance documents for the category or read-across approach for selected groups of chemicals

J.P. Rila, P.M.J. Bos, E. Hulzebos,
B.C. Hakkert

Contact: Jean-Paul Rila
Expertise Centre for Substances (SEC)
Jean-Paul.Rila@rivm.nl

This investigation has been performed by order and for the account of the Ministry of Housing, Spatial Planning and the Environment of the Netherlands, Directorate General for Environment Protection (DGM), Directorate for Chemicals, Waste, Radiation Protection (SAS), within the framework of project M/601200, Risk Assessment Tools.

RIVM, P.O. Box 1, 3720 BA Bilthoven, telephone: 31 - 30 - 274 91 11; telefax: 31 - 30 - 274 29 71

Rapport in het kort

Bruikbaarheid van richtsnoeren voor categorie- of read-acrossbenaderingen voor geselecteerde groepen van chemicaliën

De huidige internationale richtsnoeren om chemische stoffen van vergelijkbare structuur groepsgewijs te toetsen op mogelijke risico's hebben meer toelichting nodig om ze goed te kunnen toepassen. Dat blijkt uit een onderzoek van het RIVM naar de bruikbaarheid van deze richtsnoeren.

Aanleiding voor het RIVM-onderzoek is de nieuwe Europese wetgeving voor productie, handel en gebruik van chemische stoffen (REACH), die halverwege 2007 in werking treedt. Die schrijft voor dat 30.000 chemische stoffen getoetst moeten worden op mogelijke gevaren.

Om het grote aantal chemicaliën te kunnen toetsen, zijn diverse dierproefvrije methoden ontwikkeld, zoals QSARS, in vitro-methoden en de categorie- of read-acrossbenadering. Van slechts een beperkt aantal chemische stoffen is bekend welke fysisch-chemische en toxicologische kenmerken ze vertonen; denk daarbij bijvoorbeeld aan huidirritatie, oplosbaarheid in water, afbreekbaarheid in het milieu. De categorie- of read-acrossbenadering maakt gebruik van beschikbare stofinformatie om chemische stoffen met een vergelijkbare structuur waarvoor weinig van deze data beschikbaar zijn, toch te kunnen toetsen.

De huidige richtsnoeren voor deze benadering kunnen worden gebruikt als basisdocument. Enkele verbeterpunten zijn gewenst in de verdere ontwikkeling van de REACH-richtlijnen. Belangrijk aandachtspunt daarbij is een heldere definitie van de categorieën die voor read-acrossbenaderingen worden gebruikt om te voorkomen dat ongelijkwaardige data worden vergeleken. Die onderbouwing en een heldere documentatie van gegevens voor deze benadering bepalen in hoge mate de bruikbaarheid van het nieuwe systeem.

Trefwoorden: categoriebenadering, read-acrossbenadering, chemicaliën, richtsnoeren, OECD, REACH

Abstract

Workability of the guidance documents for the category or read-across approach for selected groups of chemicals

The current international guidance documents for performing a group-based assessment of the possible risks caused by chemical substances with comparable structures need further elucidation if they are to be properly used. This was the result of RIVM research on the workability of guidance documents.

This research was prompted by the upcoming European legislation on production, trade and use of chemical substances (REACH), which will come into force in mid-2007. This legislation stipulates that about 30,000 chemical substances are to be assessed on their possible risks.

Several non-animal methods such as QSARs, in vitro methods and the category or read-across approach have been developed to assess this large number of chemicals. The physico-chemical and toxicological properties are only known for a minority of all chemical substances. These may for example include skin irritation, water solubility and degradation in the environment. The category or read-across approach uses available substance information to be able to assess chemical substances with comparable structures for which only few data are available.

The current guidance document for this approach can be used as a basis, but several points still need more attention in the further development of the guidance document for REACH. One point of particular interest is establishing a clear definition of the categories for use in the read-across approach to avoid comparing unequal data. Data substantiated and clearly presented for this approach highly define the usefulness of this new system.

Key words: category approach, read-across approach, chemicals, guidance, OECD, REACH

Contents

Summary

For the majority of chemical substances there is little or no information on their hazardous properties. Within the upcoming EU regulation (REACH) for chemical substances information has to be gathered or generated for about 30,000 chemicals before the year 2015. For the reasons of animal welfare, costs and logistics it is important to limit the number of tests to be conducted. This means that suitable non-test methods have to be developed that allow regulatory decisions to be made. At present, non-animal testing techniques are seldom used to replace test data within the EU, for which there are formal data requirements in the legislation. An upcoming trend, however, is noticed in the application of the category or read-across approach, which is based on the expectation that structurally similar chemicals will have similar physical attributes and biological effects. To make a hazard assessment for the parent substance(s) and the analogues one has to judge that all or certain endpoints based on data from the parent substance(s) can cover for the analogues.

Currently there is no formal guidance document that outlines best practice for performing read-across and/or chemical grouping for the evaluation of chemical substances. However, there are guidance documents that give indications on how to do this. Therefore, the main goal of this report was to evaluate whether the OECD guidance on categories for High Production Volume chemicals and the United Kingdom Health and Safety Executive draft proposal on read-across or categories for new substances are sufficient for regulators and industries to evaluate the hazards of substances with few/no data and, if not, to provide suggestions for improvement.

The evaluation was performed by applying the above mentioned guidelines on groups of data-rich chemicals. The following groups were selected: C2-C9 backbone phthalates, C5-C9 straight-chain aliphatic hydrocarbons, butenes (C4 olefins) and a read-across example of one existing substance and two new substances. To evaluate the robustness of the example categories first the steps described in the OECD guidance were followed, which are structural similarities, applicability domain and trend analysis. Additionally, the needs and principles as defined in the UK-HSE guidance were applied. The use of alternative methods like computer programs (QSARs), analogue models and in vitro data were considered.

For the C2-C9 phthalates it was concluded that they share common functional groups but differ in the alkyl chains. Furthermore, there were no great differences in physico-chemical properties, absorption rates and metabolism. However, with respect to the endpoint reproductive toxicity the investigated phthalates did not follow a linear trend, but rather a parabolic one, which made it clear that the mechanism of action had to be elucidated. The use of alternative data like QSARs and in vitro data would not have helped in revealing this and therefore were of no use in this particular case. The C2-C9 straight-chain aliphatic

hydrocarbons also showed structural similarities and their basic biotransformation pathways were similar. However, differences were found in the induction of peripheral neuropathy upon repeated exposure. Only performance of repeated dose toxicity study, including monitoring of adequate neurological parameters, could have revealed the exceptional position of one substance in this category (n-hexane). The butenes did overall meet the steps described in the OECD chemical category concept. Alternative data, in this case QSARs, were considered valid for determining the acute aquatic toxicity, biodegradability and bioaccumulation of the substances in this category. Differences were observed for the endpoint repeated dose toxicity, for which there were studies with different timeframes leading to derivation of NOAELs based on different critical effects. Results form the read-across example showed that the use of this approach was not justified for human toxicological endpoints, because there were not enough data on identity, purity profile, functional groups and the precursors/breakdown products. Also only some physico-chemical parameters were comparable and it was not possible to interpret the differences in the endpoint repeated dose toxicity.

During the evaluation of the examples it became clear that for judging the similarities between substances in a category various types of expertise are needed and it demands a high amount of resources with respect to time. Reading across information between substances seems to work for substances that do exhibit relatively non-specific toxicological effects. For endpoints with a very narrow time window or for substances that form different (active) metabolites this approach has to be used with care. The latter one mentioned will even be more complicated by the fact that under REACH/OECD metabolic information is not required. It became also clear that base studies on which a category or read-across approach will be based, have to be of high quality. Application of category and read-across approaches needs to be justified on a case-by-case basis. Their usefulness highly depends on how well the case is substantiated and documented and particularly depends on insight into mechanisms of action.

The insights gained in this working document will be brought forward in the REACH Implementation Project 3.3 and also in the new Technical Guidance Document, which is currently under development.

Abbreviations

ADME – Adsorption Distribution Metabolisation Excretion
AIM – Analogue Identification Methodology
AR – Androgen Receptor
ATC - Technical Committee of Petroleum Additives Manufacturers
ATSDR - The Agency for Toxic Substances and Disease Registry
BBP - Benzylbutyl phthalate
CESIO - Comite Européen des agents de Surface et de leurs Intermédiaires Organiques
CNS – Central Nervous System
CONCAWE - European Oil Company Organisation for Environment, Health and Safety
DBP - Dibutyl phthalate
DEHP - Bis(2-ethylhexyl) phthalate
DEP - Diethyl phthalate
DIHP - Di-isoheptyl phthalate
DnHP - Di-n-hexyl phthalate
DnOP - Di-n-octyl phthalate
DPHP - Bis(2-propylheptyl) phthalate
DPP - Di-n-pentyl phthalate
DSL – Domestic Substances List
ECA – European Chemicals Agency
ECB – European Chemicals Bureau
EEC – European Economic Community
EFSA – European Food Safety Authority
EPA – Environmental Protection Agency
ETAD - The Ecological and Toxicological Association of Dyes and Organic Pigments
EU – European Union
FAO – Food and Agriculture Organization
GDCh - Gesellschaft Deutscher Chemiker e.V.
GHS - Globally Harmonized System of Classification and Labelling of Chemicals
HPVC – High Production Volume Chemicals
HSPA - Hydocarbon Solvents Producers Association
ICCA – International Council of Chemical Associations
Insl3 - Insulin-like hormone 3
IPCS – International Programme on Chemical Safety
IUPAC - The International Union of Pure and Applied Chemistry
JECFA - The Joint FAO/WHO Expert Committee on Food Additives
LAS – Linear Alkylbenzene Sulfonates
L(C)D50 - Dose (concentration) that kills 50% of the animals tested (LD = “lethal dose”).
LEL – Lower Explosion Limit
LOAEL – Lowest Observed Adverse Effect Level
MEHP – Mono etylhexyl phthalate
NOAEL(C) – No Observed Adverse Effect Level (Concentration)
NONS – The Notification Of New Substances
NTP – National Toxicology Program
OECD – Organisation for Economic Co-operation and Development
OPPT – Office of Pollution Prevention and Toxics
PAE – Phthalic Acid Ester

PMN – Premanufacture Notice
QSAR – Quantitative Structure-Activity Relationship
RAR – Risk Assessment Report
REACH – Registration Evaluation and Authorisation of Chemicals
RIP – REACH Implementation Program
SAR – Structure-Activity Relationship
SCF – Scientific Committee on Food
SIAM – SIDS Initial Assessment Meeting
SIAP – SIDS Initial Assessment Profile
SIAR – SIDS Initial Assessment Report
SIDS – Screening Information Dataset
T - Testosterone
TAPIR – RIP 3.3 A Project on Information Requirement for REACH
TLV – Threshold Limit Value
TPA – Tonnes Per Annum
TSCA – Toxic Substances Control Act
UK-HSE – United Kingdom Health and Safety Executive
WHO – World Health Organization

1. Introduction

Currently, there are several regulatory programs that are designed to provide available, or generate new, hazard information on chemical substances. The OECD (Organisation for Economic Co-operation and Development) High Production Volume Chemicals (HPVC) program has listed 5000 substances in 2004, which will be/have been subject of an evaluation of their hazardous properties. In Europe, Directive 67/548/EEC requires new substances to be tested and assessed for possible risks to human health and the environment before they are marketed in volumes starting at 10 kg/year. Within the EU Existing Substances program, which is driven by prioritisation, (draft) risk assessment reports have been finished for about 300 chemicals, which means that in the EU information for the major part of HPV substances is lacking. Recently the European Commission published a proposal for a new chemicals legislation, REACH, in which an approach to develop necessary information for all chemical substances with a tonnage level of over >1 tpa (tonnes per annum) is outlined. The consequence of this is that information has to be gathered or generated for about 30,000 chemicals before the year 2015.

The goal of this report will be to evaluate whether the OECD guidance on categories for HPV chemicals (OECD, 2005) and the UK-HSE (United Kingdom Health and Safety Executive) draft proposal on read-across for new substances (Hanway, 2002a,b) are sufficient for regulators and industries to evaluate the hazards of substances with few/no data and, if not, to provide suggestions for improvement. In other words, do these guidelines provide enough tools/methods for doing this evaluation in such a way that the safe use will not be compromised? To achieve this goal, it was the intention to select examples of categories containing data-rich substances and apply the above mentioned guidance/proposal on these. However, as it is the purpose of a category or read-across approach to fill in data gaps for certain substances these examples were not easily found. Therefore, also examples on experiences from the evaluation of OECD categories and New Substances will be described. Insight/experience will be gained on the difficulties that can occur when evaluating a category or read-across approach. The chemical category concept as defined by the OECD (subsection 3.2.2 steps 6 up to 13, 2005) will be followed to see whether we can elucidate the mechanism of action, establish the parameters that need to be looked at, derive a critical effect level and label substances. In case that these aspects cannot be resolved with the existing data the use of alternative methods like computer programs (QSARs), analogue models and in vitro data will be considered. This will be discussed in the light of the data requirements for the specific frameworks. The outcome of this exercise is intended to provide input in the OECD HPV Framework, the New Substances Framework as well as in the REACH Implementation Projects (RIPs).

Within all these programs there is understanding that for reasons of animal welfare, costs and logistics, it is important to limit the number of tests to be conducted. It is also recognised that this means that suitable non-test methods have to be developed that allow regulatory decisions to be made. Non-animal testing information could either be used as supplementary information, to be used in a weight of evidence approach, or as stand-alone. Use of such data contributes to the reduction of animal use/testing.

The principle of development and use of non-animal testing methods is based on the expectation that structurally similar chemicals will have similar physical attributes and biological effects. This underlying premise of similarity can be used in hazard (and risk) assessment when there are inadequate test data to estimate missing values.

At present, these non-animal testing techniques are seldom used to replace test data within the EU, for which there are formal data requirements in the legislation. In order to use these techniques with the same confidence as animal test data, a number of issues needs to be addressed. This includes the way the surrogate data are obtained (use of QSARs, test batteries et cetera), the number and type of endpoints that need to be addressed in the specific framework, how the data are interpreted in relation to criteria, and the confidence placed in the models used (validity).

To make a hazard assessment for the parent substance(s) and the analogues a regulatory toxicologist has to judge that all or certain endpoints based on data from the parent substance(s) can cover for the analogues. There is currently no formal guidance documentation that outlines best practice for performing read-across and/or chemical grouping for the evaluation of chemical substances.

This report will focus on the current understanding of the category or read-across approach as described in the before mentioned OECD guideline and UK-HSE proposal. After a description of the methodology used for this report (chapter 2), the OECD guidance, the UK-HSE approach and guidance/information from other frameworks (EPA, Canada, Flavourings) will be described in chapter 3. This chapter will give a literal description of these guidelines and is largely based on work that was already carried out under the REACH Implementation Program 3.3 on non-testing considerations (TAPIR, 2005). The TAPIR report gives an overview of current guidance/practice, describes on-going development and identifies further needs in relation to the information requirements according to the REACH proposal. The OECD guidance for categories and the UK-HSE proposal for read-across will then be applied on selected examples (phthalates, butenes, aliphatic hydrocarbons and read-across for some new substances) in chapter 4. Furthermore, this chapter will give an analysis of the selected samples and describe the problems encountered in the current approach from the viewpoint of a regulatory toxicologist. Aspects that will be addressed are metabolism and kinetics, which are no part of the data requirements for industrial chemicals in general. Chapter 5 will then provide a discussion, conclusions and recommendations on the evaluation of the category or

read-across approach and will give suggestions for improvement of current guidance/approaches.

2. Methods

This chapter will address the methodology used for this report. The report of RIP 3.3 Non-testing considerations (2005) was used as starting point, as the current practice and limitations of available guidance in different frameworks were already pointed out in that document. Existing guidance on chemical categories and read-across, as described in the OECD HPV guidance (OECD chapter 3.2, 2005) and the draft UK-HSE approach for New Substances (Hanway, 2002a,b) are summarized in chapter 3. The full text of these documents is provided in Appendices 1, 2 and 3. Guidance/information from other frameworks was described in short in chapter 3.

Groups of data-rich chemicals (existing and non-existing categories) were searched for. The following categories will be evaluated in this report:
- C2-C9 backbone phthalates;
- C5-C9 straight-chain aliphatic hydrocarbons;
- butenes (C4 olefins);
- read-across of 1 existing substance and 2 new substances.

The motivation for the selection for each of these examples will be described in chapter 4.

In the description and analysis of the selected examples on categories the following steps will be taken:

1. Available information on structures, identity, physico-chemical properties and (eco)toxicological data will be summarized in a two-dimensional matrix (see Figure 2.1). Different category members occupy different columns, and the different category endpoints occupy different rows.
2. The selected examples will be evaluated using the chemical category concept as defined by the OECD (see Appendix 1, steps 6 up to 13). These are:
   - Structural similarities in groups of chemicals can result in physico-chemical and toxicological properties that are likely to be similar or follow a regular pattern. These structural similarities may create a predictable pattern in any or all of the following parameters: physico-chemical properties, environmental fate and environmental effects, and human health effects. The similarities may be based on the following:
     - a common functional group;
     - likelihood of common precursors and/or breakdown products, via physical or biological processes, which result in structurally similar chemicals;
     - an incremental and constant change across the category.
   - The applicability domain. This domain identifies the physico-chemical property space within which the chemical category is considered reliable. In the context of a chemical category, it can be considered to identify the ranges of physico-chemical, environmental, and (eco)toxicological properties within which reliable estimations can be made of missing data points, by the use of trend analysis, read-across and

(Q)SARs and or AARs (see chapter 3 for definitions). It can also be considered as a set of inclusion and/or exclusion rules that identifies the ranges of values within which reliable estimations can be made for category members.

- Trend analysis will be applied on the category under evaluation when the members of a category exhibit a series of increasing or decreasing values for a given endpoint. Interpolation is the estimation of a value for a member using measured values from other members on "both sides" of that member within the defined category spectrum (see Figure 2.1), 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 (see Figure 2.1).

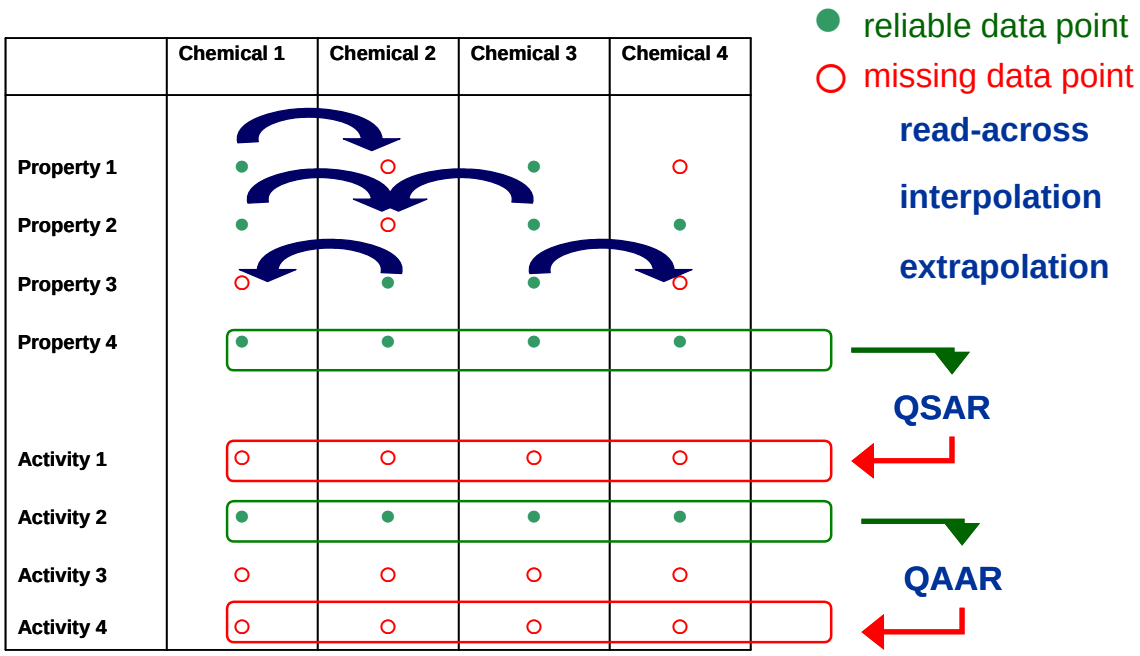


Figure 2.1 Graphical representation of a chemical category and ways of filling in data gaps (QSAR: Quantitative Structure-Activity Relationship; QAAR: Quantitative Activity-Activity Relationship)

The possibilities of identifying similarities and differences between the substances and grouping them will be discussed. If applicable, the needs and principles (see Appendix 2) as defined by the UK-HSE will also be applied.

The use of (Q)SAR models (METEOR, DEREK, ECOSAR/EPIWIN) and in vitro methods within these approaches will be discussed. If identified, suggestions for improvement of current guidance will be made.

3. Existing guidance on chemical categories and read-across

In this chapter an overview will be given on existing guidance/information on categories and read-across. First in section 3.1 proposed working definitions from TAPIR will be given for the most commonly used terms when looking at categories or read-across. In section 3.2 available guidance, current practice and on-going development on categories and read-across as described in TAPIR (Non-testing considerations report of the Information Working Group 3 within the RIP 3.3-1 consortia) will be given. The mentioned guidance on categories of the OECD and on read-across of the UK-HSE will provide the background for the evaluation of the examples in chapter 4. A description of guidance in other frameworks will be provided to give an overview of the current status. The full text of the guidance/proposal on categories (OECD HPVC), read-across for human toxicological and ecotoxicological endpoints (UK-HSE New Substances approach) is given in the Appendices 1, 2 and 3, respectively.

3.1 Definitions

In TAPIR working definitions are proposed for read-across (qualitative/quantitative), chemical categories and (Q)SARs. Most of the definitions are taken directly from the OECD guidance on categories.

Chemical categories
For chemical categories the same definition is proposed as used in the OECD guidance document (see Appendix 1 point 6).

Read-across
The process by which one or more properties of a given chemical are inferred by comparison of that chemical with a chemical(s) of similar molecular structure(s) and physico-chemical properties, for which the properties of interest are known. This approach can be used to assess physico-chemical properties, toxicity, environmental fate and ecotoxicity. Read-across can be performed qualitatively or quantitatively (see for the definitions of these latter terms Appendix 1 point 9).

(Q)SARs
A (Q)SAR consists of a relationship between the chemical structure, or physico-chemical representations thereof and the outcome in a test for an endpoint (biological or other physico-chemical properties). Two types can be distinguished:

1. SARs are qualitative relationships in the form of structural alerts that incorporate molecular substructures of fragments related to the presence or absence of activity

2. QSARs are quantitative models yielding a continuous or categorical result

Within the OECD guidance on chemical categories read-across, SARs and QSARs are mentioned as procedures to fill data gaps within categories. From the definitions above the main distinction that can be made is that read-across/closest analogue analysis will normally be performed between one data-rich substance and a substance for which limited data are available. In the category approach the similarity of a pattern for several chemicals will be evaluated. Read-across can be one tool to do this, but interpolation and extrapolation and (Q)SARs can also be considered in defining the applicability domain of a chemical category.

Within a category different members can be selected for the endpoint desired - i.e., those selected for a category approach for environmental effects endpoints may not be suitable for assessing human health effect endpoints. Thus if the available test results show that the chemicals in a category behave in a similar or predictable manner, then interpolation and/or extrapolation may be used to assess the chemicals in lieu of conducting additional testing. This may be done in one of two ways:

- using data on individual substances (components) to derive an understanding of the properties on the non-tested members of the category, or
- using data on the individual components within the complex substance, e.g. use of the hydrocarbon block method or specific chemical marker/surrogates to derive an understanding of the properties of the group.

3.2 Current status on the use of categories and read-across

3.2.1 Chemical categories

3.2.1.1 Available guidance

OECD

Guidance on the formation and use for chemical categories for fulfilling data requirements has been published by the OECD as part of the OECD Manual for Investigation for HPV Chemicals (OECD, 2005). This guidance is used among others for fulfilling the data requirements within the OECD HPV Chemicals Programme. The same guidance document is published by the US-EPA for use within the US HPV Challenge Programme. The OECD guidance document is currently being revised and goes together with the draft guidance on the grouping of chemicals, which is set up within RIP 3.3-2 task 3 for implementation within REACH. For a full text see Appendix 1.

The draft guidance document addresses the following issues:

– definitions and explanations of the chemical category concept
– general approach for developing categories
– differences in grouping for different endpoints
– use of QSARs for the development of a category

– guidance on different types of categories (i.e. chain-length, metabolic pathways, isomers and their mixtures, complex substances, metal and metal compounds)

Canada

Environment Canada also uses an analogue approach and the following general rules of thumb (but recognizes that there will always be exceptions):

- An analogue should preferably contain most, if not all, of the same structural features as the substance of interest
- An analogue should have approximately the same molecular weight as the substance
- An analogue should have water solubility similar to that of the substance of interest
- For persistence, an analogue should have the same reactivity or stability as the substance of interest
- For an endpoint of interest, the relevant molecular descriptors of an analogue should be of comparable value to those of the substance

It is recognized that different analogues may be selected for different endpoints, e.g., an analogue selected for a P endpoint may not be suitable for determining a B endpoint. Environment Canada and Health Canada rely on many of the on-line databases, but have also created extensive in-house databases for physico-chemical properties and toxicity that are searchable by structure using the Chemfinder software (http://chemfinder.cambridgesoft.com) or ISISBase (http://www.mdli.com/products/framework/isis_base/index.jsp). Most property and toxicity data for new substances are stored in these databases and a large analogue database has been created by Environment Canada for DSL (Domestic Substances List) Categorisation. Sources for searches of analogue(s) may be publicly available databases that allow substructure searches, commercial databases, proprietary databases and books and reference sources. However, searching for analogues from books is resource demanding and difficult to perform systematically, as compared to electronic substructure searching, which for instance can be performed on skeletal framework, functional groups and substituents.

Methods for analogue selections are expert knowledge in combination with electronic substructure searching and automatic tools using molecular similarity indexes (e.g. the Tanimoto similarity index).
The pharmaceutical industry, which are the predominant users of the concept of molecular similarity, are employing similarity methods in a wide range of applications e.g. virtual screening, estimation of absorption, distribution, metabolism, excretion and toxicity (ADME/Tox) and prediction of physico-chemical properties (solubility, partioning, et cetera).

Industry

CONCAWE has developed the hydrocarbon block method, and presented in 2004 the gasoline risk assessment based on this approach. A generic approach (primarily for evaluating environmental hazard) is now being developed for higher boiling point, complex hydrocarbon substances, which develops further the concept of hydrocarbon blocks and how their use can be extended to a broader range of multi-component complex chemicals.

A number of industry sectors have applied the principles of "grouping" for use in evaluation of health and environmental hazard properties. Examples, including rationales for grouping, include petroleum substances (CONCAWE, 1998), dyes and pigments (ETAD, 2001), surfactants (CESIO, 2000, 2003) hydrocarbon solvents (HSPA, 2002), acrylate resins (UV/EB Acrylate Resins, 2003), petroleum additives (ATC, 2000) and bitumen (Eurobitume, 2002).

3.2.1.2 Current and past practice with categories

Grouping has also been used tentatively within the EU Notification scheme for new chemicals. As an example, a request received by the UK Health and Safety Executive (UK-HSE) involved a series of four structurally similar substances differing in their numbers of carbon atoms. The result was a full base-set testing of the lowest homologue of the series, and a limited testing on the highest homologue. For the other group members, all toxicological endpoints used for base-set notification were filled by read-across. Based on experience in the use of read-across data under NONS the UK-HSE (Hanway, 2002a,b) has developed a strategy that to their opinion can be used when considering whether the use of toxicological read-across data is scientifically justified. The strategy incorporates a series of needs and principles that have emerged during their evaluation of read-across data and has been considered a useful tool in assessing whether a read-across argument is valid. Strategies for human toxicological and environmental endpoints were developed separately. For a full text see Appendix 2.

The principle of "grouping" of chemicals has been applied under various Community legislative provisions for both hazard identification and risk assessment. Annex I to Directive 67/548 EC contains a significant number of "group" entries, and groups of petroleum and coal derived substances have been reviewed as groups but are listed in Annex I on a substance by substance basis. The Annex itself is built on a simplified category approach, as Annex I entries are grouped by Index Number into categories of inorganic compounds based on atomic number or into 20 different categories of organic compounds. Risk assessments undertaken under Regulation 793/93 EC have also "grouped" chemicals for evaluation, particularly metal compounds. Annex 1 to Regulation 793/93 EC formally recognised grouping of petroleum substances for both registration and risk evaluation purposes and CONCAWE are currently producing voluntary risk assessments for groups of petroleum products using this approach. The process for establishing occupational exposure

limits at National and Community level has used a “grouping” approach where appropriate and other bodies, such as WHO (IPCS and IARC) have also used “grouping” principles from time to time.

A general approach to chemical categories has long been practised in e.g. the pharmaceutical industry. A similar approach has also been considered for pesticides, biocides and to some degree on certain types of industrial chemicals. A Nordic strategy has been developed (Wold, 1987) and an attempt made to apply this in practice (Jensen et al., 1989). This approach has been used by other groups (Cesareo et al., 1987). Attempts have also been made to cluster compounds as a starting point for a systematic approach (Nouwen and Hansen, 1994).

For a safety evaluation of chemically defined flavouring substances a stepwise approach, the procedure, is used, which integrates information on intake from intended use in foods, structure activity relationships, metabolism and, when needed, toxicity (EC, 2000). One of the key elements in the procedure is the subdivision of flavourings into three structural classes (I, II, III) for which thresholds of concern (human exposure thresholds) that are not considered to present a safety concern have been specified.
In addition to the data provided for the flavouring substances to be evaluated (candidate substances), toxicological background information available for compounds structurally related to the candidate substances is considered (supporting substances), in order to assure that these data are consistent with the results obtained after application of the procedure. The procedure is not to be applied to flavourings with existing unresolved problems of toxicity e.g. when a substance is suspected to possess genotoxic activity. Therefore, the right is reserved to use alternative approaches if data on specific flavourings warranted such actions.

3.2.1.3 On-going further development of existing guidance on categories

The guidance document for the development and use of chemical categories within the OECD HPV Chemicals Programme is currently being revised and goes together with the draft guidance on the grouping of chemicals, which is set up within RIP 3.3-2 task 3 for implementation within REACH. In parallel to the guidance document on the formation and use of chemical categories, the development of a (Q)SAR Application Toolbox is foreseen within the OECD work programme on (Q)SARs. The (Q)SAR Application Toolbox will contain a module aimed at facilitating the development of categories for organic chemical with multiple functional groups. The Toolbox will also identify chemicals which might be placed in a particular category but which might have significant metabolic pathways from the other members. These new approaches should be evaluated over the next three years.

3.2.2 Read-across

3.2.2.1 Available guidance

Guidance on the use of read-across is available for the US EPA High Production Volume (HPV) Challenge Program (http://www.epa.gov/cgi-bin/epaprintonly.cgi). The same guidance is also included in the OECD "Manual for Investigation of HPV Chemical" (http://www.oecd.org/document/7/0,2340,en_2649_201185_1947463_1_1_1_1,00.html). Further guidance is available for the Canadian New Substance Program (http://www.ec.gc.ca/substances/nsb/cpguide/eng/cp_s5_e.htm)

The pharmaceutical industry, which is the predominant user of the concept of molecular similarity, is employing similarity methods in a wide range of applications e.g. virtual screening, estimation of absorption, distribution, distribution, metabolism, excretion and toxicity (ADME/Tox) and prediction of physico-chemical properties (solubility, portioning, et cetera).

3.2.2.2 Current practice

The U.S. EPA within their PMN-programme, especially for human health hazard assessments, is extensively using closest analogue analysis.
The concept of analogues is also widely used in the OECD HPV Chemicals Programme. No statistics are available regarding the frequency of use of the concept in initial hazard assessments agreed within this programme.

3.2.2.3 On-going development

RIP 3.3-2 task 3 in collaboration with OECD is revising the guidance on categories and read-across.

The U.S. EPA/OPPT is developing software, the Analogue Identification Methodology (AIM), which presently is in beta testing. AIM will be a web-based, computerised tool that identifies chemical analogues based on structure. It will identify closely related structures for which data are available, and will point the user to specific, publicly available data sets or other sources of information where data can be obtained.

4. Results and analysis

In the following paragraphs examples of data rich substances are described, for which the category or read-across approach was applied. The earlier mentioned guidance/proposal on categories of the OECD and on read-across of the UK-HSE will be applied on the examples to see whether these guidelines provide a systemic approach for evaluation. The chemical category concept as defined by the OECD (see chapter 2 Methods) will be followed to see whether we can elucidate the mechanism of action, establish the parameters that need to be looked at, derive a critical effect level and label substances. The needs and principles as defined by the UK-HSE will also be used on these examples if applicable. For defining the applicability domain of the selected categories the use of alternative methods like computer programs (QSARs), analogue models and in vitro data will be considered.

4.1 Category C2-C9 backbone phthalate esters

4.1.1 Introduction

Based on a literature search on data rich phthalates and work that was performed by Fabjan et al. (2006) a group of 9 ortho phthalates (with backbone C2-C9) were selected. In the present context “backbone” will be regarded as the longest straight alkyl chain and not the total number of carbon atoms in a side chain. The selection was made on the basis of the amount of data available (all these phthalates have been evaluated in (inter)national frameworks) and the concern for reproductive effects for phthalates which have the ester in the ortho position and side chains of length C4 to C6. Fabjan et al. (2006) looked in more detail at the mechanistic background of the reproductive effects of ten ortho phthalates. Five of the phthalates in this category have already been or are under evaluation in the EU Existing Substances framework and/or the OECD HPV program: DBP (ECB, 2004a; OECD, 2001), BBP (ECB, 2004c), DEHP (ECB, 2001; OECD, 2005a), DINP (ECB, 2003; OECD, 1999a) and DIDP (ECB, 2003a; OECD, 1999b). For all of these substances specific studies (i.e. at least a two-generation study with rats) are available to cover the endpoint reproductive toxicity.
For the other four no two-generation studies with rats were available. The conclusion on the endpoint reproductive toxicity was based on other studies (repeated dose toxicity studies, continuous breeding studies). DPHP (backbone C7 branched with a propyl side chain) was evaluated within the category of High Molecular Weight Phthalates with backbone C7 to C12 (OECD, 2004a). DEP was evaluated by IPCS (2003), DnHP and DnOP by NTP (2003a and 2003b).

DEP, DBP, BBP, DnHP, DEHP, DPHP and DnOP are “well defined phthalates”, DINP and DIDP are complex mixtures (ECB, 2003; ECB, 2003a). In case of DINP, in the EU Risk Assessment Report, three different DINPs were identified under two CAS numbers and two

structural formulas were recovered under these two CAS numbers. In the case of DIDP two different DIDPs were identified under two CAS numbers, with respective structural formulas (ECB, 2003; ECB, 2003a).

4.1.2 Use of phthalate esters

Diesters of phthalic acid with straight or branched chain alcohols (PAEs), commonly known as phthalates, are ubiquitous industrial chemicals which are among others used in consumer products (soaps, shampoos, and other cosmetics), plastics (including food packaging), paints, enteric coatings in some medications, and pesticide formulations. As they are not covalently bound to the polymer, they are fairly easily released to air, water, saliva, blood, food and other extracting material.

4.1.3 Category evaluation

Following the OECD guidance on chemical categories, substances that are to be grouped will have physico-chemical and toxicological properties that are likely to be similar or follow a regular pattern as a result of structural similarity. In the following the steps as described in chapter 2 (Methods) will be evaluated.

4.1.3.1 Graphical presentation as a two-dimensional matrix of the C2-C9 phthalates category

In Table 4.1 the physico-chemical properties of the substances in this category are shown. Furthermore, Table 4.2 gives an overview of the most relevant data for the C2-C9 phthalates category for reproductive toxicity.

Table 4.1 Physico-chemical properties of C2-C9 backbone phthalate esters

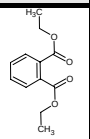
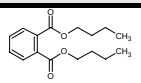
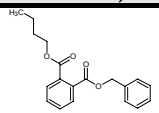
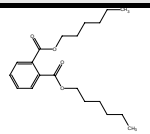
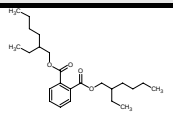
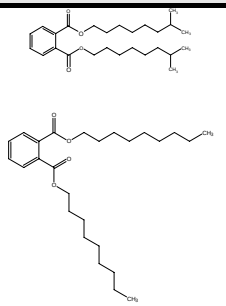
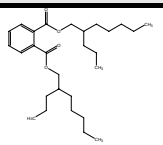
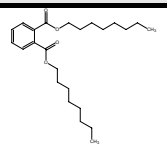
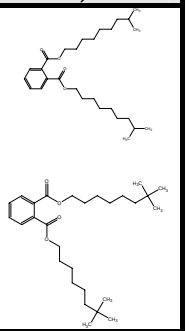
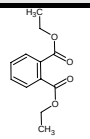
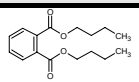
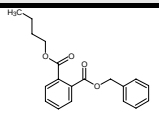
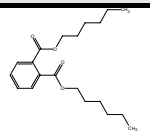
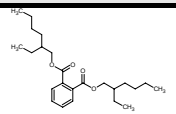
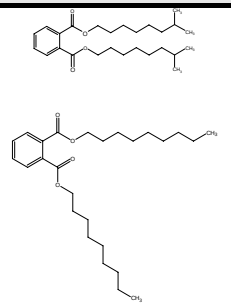
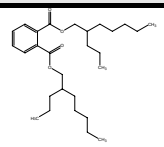
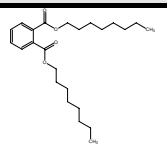
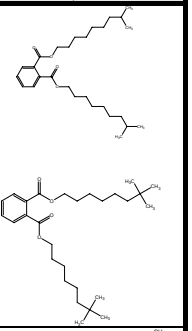
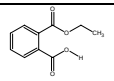
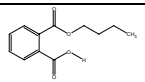
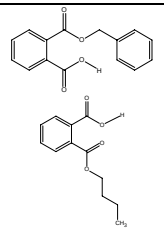
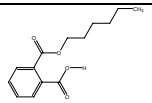
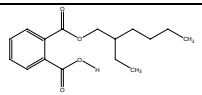
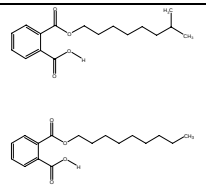
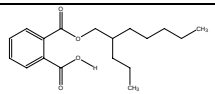
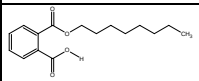
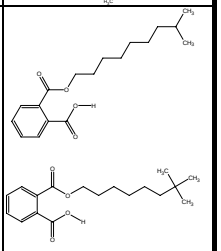
	DEP 84-66-2 C2	**DBP 84-74-2 C4**	**BBP 85-68-7 C4 (straight chain) C7 (aromatic chain)**	**DnHP 84-75-3 C6**	**DEHP 117-81-7 C6 (ethyl side chain)**	**DINP 28553-12-0 68515-48-0 C6-C7-C8-C9 (branched or straight)**	**DPHP 53306-54-0 C7 (propyl side chain)**	**DnOP 117-84-0 C8**	**DIDP 26761-40-0 68515-49-1 C7-C8-C9 (n methyl side chains)**
Structural formula									
Molecular weight	222.30	278.34	312.35	334.4	390.6	average 420.6	446.7	390.4	average 446.7
State of substance at 20ºC	liquid	liquid	liquid	liquid	liquid	liquid	liquid	liquid	liquid
Melting point [ºC]	-40.5	-69	< -35	-58	-55	-50	-48	-25	-45
Boiling point [ºC]	295	340 at 101 kPa	370 at 1 kPa	210 at 0.6 kPa	230	250 at 7 hPa	254	230 at 6 hPa	267 at 5 hPa
Vapour pressure [Pa]	0.046	0.016	0.00004	0.002	0.000034	6E-05	0.0000037	0.019	2.8E-05
Water solubility [mg/l]	1000	10	2.8	0.05	0.003	0.0004	0.0001	0.0005	0.0002
Log Kow	2.47-2.51	4.75	4.84	6.3	7.5	8.8	> 6.0	8.6	8.8

Table 4.2 Data availability on reproductive toxicity for C2-C9 backbone phthalate esters - the different studies are indicated by numbers

	DEP **84-66-2** **C2**	**DBP** **84-74-2** **C4**	**BBP** **85-68-7** **C4 (straight)** **C7 (aromatic)**	**DnHP** **84-75-3** **C6**	**DEHP** **117-81-7** **C6 (ethyl side chain)**	**DINP** **28553-12-0** **68515-48-0** **C6-C7-C8-C9** **(branched or straight)**	**DPHP** **53306-54-0** **C7 (propyl side chain)**	**DnOP** **117-84-0** **C8**	**DIDP** **26761-40-0** **68515-49-1** **C7-C8-C9** **(n methyl side chains)**
Structural formula									
Metabolite(s)									
Oral absorption	> 90% (rats)	≥ 90% (rat and hamster)	~ 80%	no data	> 50% (rat) saturable mechanism could be as high as 75% (single dose) 90% (repeated dose)	> 50% (rat) saturable mechanism could be as high as 75% (single dose) 90% (repeated dose)	no data	at least 31% (rats)	decreases with increasing dose: 56% at low dose of 0.1 mg/kg 17% at high dose of 1000 mg/kg
Study	[1]diet study [2]continuous breeding study	[1]two-generation study [2]continuous breeding study	two-generation study	[1]oral gavage studies [2]continuous breeding protocol test	[1]three-generation study diet [2]continuous breeding study	[1]two year diet study [2]two-generation study	[1]oral diet study [2]developmental screening [3]developmental	continuous breeding study	two-generation studies
Species	[1]rats [2]mice	[1]rats [2]mice	rats	[1]rats [2]mice	[1]rats [2]mice	[1]mice [2]rats	rats	mice	rats

	DEP 84-66-2 C2	**DBP 84-74-2 C4**	**BBP 85-68-7 C4 (straight) C7 (aromatic)**	**DnHP 84-75-3 C6**	**DEHP 117-81-7 C6 (ethyl side chain)**	**DINP 28553-12-0 68515-48-0 C6-C7-C8-C9 (branched or straight)**	**DPHP 53306-54-0 C7 (propyl side chain)**	**DnOP 117-84-0 C8**	**DIDP 26761-40-0 68515-49-1 C7-C8-C9 (n methyl side chains)**
Exposure time	[1]1 week	[1]119 days [2]115 days		[1]4 days			[1]90 days [2]gavage day 6 to 15 [3]gavage days 6 to 19		
Doses	[1]2000 mg/kg bw/day [2]3640 mg/kg bw/day	[1]0, 52, 256 and 509 mg/kg bw for males and 0, 80, 385 and 794 mg/kg bw for females [2]0, 40, 420 and 1410 mg/kg bw/day	0, 20, 100 and 500 mg/kg bw/day	[1]2400 mg/kg bw/day and 1824 mg/kg bw/day [2]380, 800 and 1670 mg/kg bw/day	[1]0.1, 0.5, 1.4, 4.8, 14, 46, 359, and 775 mg/kg/day [2]20, 200 and 600 mg/kg bw/day	[1]0, 500, 1500, 4000 or 8000 ppm in diet [2]125, 250, 500 (i.e. 0.2, 0.4, or 0.8% in diet)	[1]50, 250, and 1500 mg/kg bw/day [2]20, 200, 1000 mg/kg bw/day [3]40, 2000 and 1000 mg/kg bw/day	7500 mg/kg feed	[1]150, 300,600 (ie. 0.2, 0.4, and 0.8% in diet) [2]17, 50, 150, 300 (ie. 0.02, 0.06, 0.2, and 0.4% in diet)
Critical effect	[1]decreased testosterone concentration in testes and serum (approximately 40%), ultrastructural changes in Leydig cells (for 2 days) [2]decreased epididymal sperm concentration (F1) and number of live F2 pups	[1]embryotoxicity and testicular effects [2]clear effects on female fertility at highest dose	male and female reproductive organs (ovaries, testis)	[1]testicular atrophy at high doses, no effects on testis at lower dose [2]significant reduction of fertility in males and females and effects on testis at highest dose	[1]dose-dependent effects on numerous testis-related parameters [2]significant dose dependant decrease in fertility, effects on testis at highest dose	[1]decreased absolute and relative (to brain weight) testis weight [2]No effect on fertility No effect in females No effect in males	[1]no effects on reproductive organs [2]no –treatment-related findings [3] pups slightly increased rates of soft tissue, skeletal, and total variations. dams decreased food consumption, body weight loss and toxicity associated early resorptions.	decrease of relative seminal vesicles weight	[1]no effect on fertility no effect in females no effect in males [2]maternal body weight gain, food intake suppressed; reduction in offspring survival and body weights
Critical effect level [mg/kg bw/day]		[1]LOAEL 52 (embryotoxicity) NOAEL 385 (maternal tox) [2]NOAEL 420 (parental and embryotoxicity)	NOAEL 100		[1]NOAEL 4.8 (testicular tox and dev tox) 46 (fertility) [2]NOAEL 20	[1]1500 ppm (276 mg/kg/d for reproductive effects) [2]NOAEL 500 (0.8%)	[2]NOAEL 1000 [3]NOAEL 200 (maternal and dev)		[1]NOAEL 600 (0.8%) [2]NOAEL 300 (0.4%, reproductive effects) NOAEL 50 (0.06%, maternal toxicity)
Classification and labelling Annex I (ECB, 2004b)	no entry	R61, R62	R61, R62	no entry	R60, R61	no entry	no entry	no entry	no entry
Sources	IPCS (2003) GDCh (1992) ATSDR (1995)	ECB (2004a) OECD (2001)	ECB (2004c)	NTP (2003a)	ECB (2001) OECD (2005a)	ECB (2003), OECD (1999a)	OECD (2004a)	NTP (2003b) ATSDR (1997)	ECB (2003a), OECD (1999b)

4.1.3.2 The OECD chemical category concept

The similarities in physico-chemical properties, environmental fate and toxicology of a group of chemicals may be based on three bullet points (see point 6 of Appendix 1). In the following these points are applied on the category under investigation.

Common functional groups

For the evaluation of this bullet identity, functional groups and the purity profile will be considered.

Identity

Phthalate esters are the dialkyl or alkyl aryl esters of 1,2-benzenedicarboxylic acid. They are synthesized from phthalic anhydride and the appropriate alcohol. In general, the term of phthalate esters is used to address the ortho form of benzedicarboxylic acid. In generic use the ortho configuration is implied, while the meta and para structural configuration are known as isophthalates and terephthalates, respectively. PAEs (phthalic acid esters) are generally colourless liquids and most of them are poorly soluble in water (Thomas and Thomas, 1984).

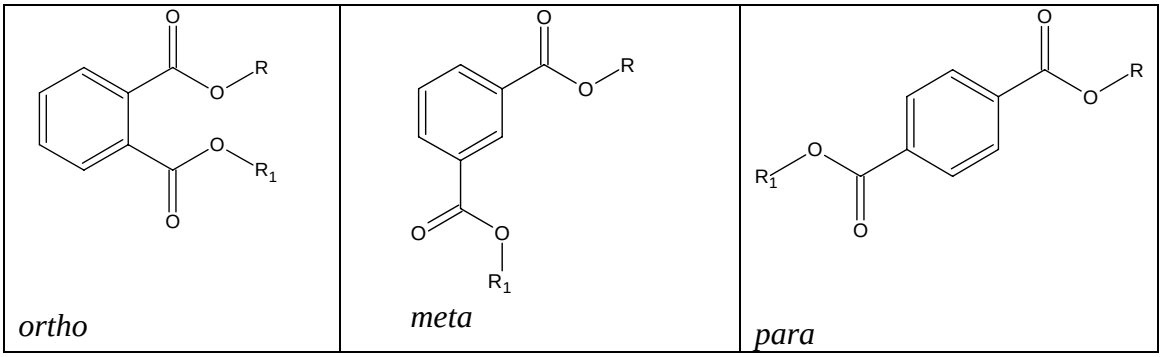


Figure 4.1 General structural formula for phthalates with esters on ortho, meta and para positions.

Functional groups

The C2-C9 backbone phthalate ester category consists of esters at the ortho position with an alkyl carbon backbone (branched, linear or aromatic (BBP)) from 2 to 9 carbon (C) atoms (see also Table 4.1). Substances in this category have the following basic structure with alkyl groups (R_1) of different lengths and or structure (branched, straight, aromatic):

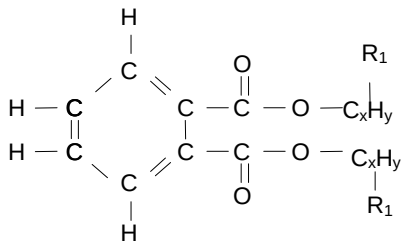


Figure 4.2 Basic structure of phthalate esters

Functional groups that can be identified are the diester structure identical for all members of this category and the alkyl chain, which is different for the individual members of this category. Furthermore, DINP and DIDP are mixtures of different backbone phthalates. The molecular structures for both substances as presented in Table 4.1 are recovered when searching under the used CAS numbers. The correct structures can only be estimated. Using data on the repartition of alcohols used for the manufacture of the DINPs, ECPI (1997) has made an estimation of the different chain structures that may be present in DINPs (Table 4.3). Based on nonene (CAS 97593-01-6) isomer distribution analysis and 1H-NMR analysis of isodecyl alcohol, an estimation of key isomeric structures of isodecylalcohol - hence of DIDP were provided by ECPI (1998, personal communication) in Table 4.4.

Table 4.3 Best estimate in content (%) of the different chain structures of the DINPs

	DINP 1	DINP 2	DINP 3
Methyl ethyl hexanols	5 - 10	5 - 10	65 - 70
Dimethyl heptanols	45 - 55	40 - 45	20 - 25
Methyl octanols	5 - 20	35 - 40	-
n-Nonanol	0 - 1	0 - 10	-
Isodecanol	15 - 25	-	-

Table 4.4 Best estimates of the different chemical structures of DIDP

Longest chain (estimates)	DIDP (CAS 68515-49-1 & CAS 26761-40-0)	Best estimated content (%)
C7	tri-methyl heptanols	0-10
C8	di-methyl octanols	70-80
C9	methyl nonanols	0-10
C10	n-decanol	0

Looking at these tables it can be concluded that depending on the kind of DINP the backbones can consist of C6 (branched with methyl and ethyl side chain), C7 (branched with two methyl side chains), C8 (branched with a methyl side chain) and C9 (straight). DIDP consists mainly of phthalates with backbone C8 (branched with two methyl side chains).

Purity profile

In addition, the UK-HSE (Hanway, 2002a,b) also pointed out that the purity and impurity profiles of substances, for which read-across is going to be applied, should be assessed. All "well-defined" phthalates in the C2-C9 backbone phthalate ester group have a very high purity (> 97%). The mixtures DINP and DIDP have a high degree of purity in terms of ester content. The content (%) of the different strain structures in DINP and DIDP is different.

Likelihood of common precursors and/or breakdown products, via physical or biological processes, which result in structurally similar chemicals.

First, it has to be noted that in order to evaluate this likelihood, data on metabolism and toxicokinetics are needed. Typically, these types of studies are not regulatory required and therefore would require a sponsor of the chemical to do additional work beyond what is normally considered necessary.

However, for the C2-C9 backbone phthalate esters there is extensive data on the kinetics and metabolism. Oral absorption is relatively high for all the substances, but appears to be a saturable process for phthalates with side chains of C6 or longer. It is assumed, because of the rapid hydrolysis process after oral ingestion, that these substances are mainly absorbed as monoester derivatives rather than as diesters (Fabjan et al., 2006).
It appears that for all the investigated substances the first step in metabolism is rapid hydrolysis to monoester (see also Table 4.2), which can be then followed by further hydrolysis and/or oxidation and glucuronidation. One exception is BBP, which is partially hydrolyzed primarily to MBuP and benzyl alcohol with monobenzyl phthalate and n-butanol as minor products of hydrolysis. There is a preference for hydrolysis of benzyl ester, resulting in a preponderance (approximately 3:1) of monobutyl phthalate in the urine. All of the substances are considered to be eliminated from the organism in a few days and none of them is considered to accumulate in the organism (Fabjan et al., 2006).

An incremental and constant change across the category (e.g. a chain-length category, see Appendix 1 point 26).

This approach can be evaluated in two steps: first the physico-chemical properties followed by evaluation of reproductive toxicity.

Physico-chemical properties:
Molecular weights of the substances range from approximately 222 for DEP to approximately 447 for DIDP. Most of them have low water solubility, with exception of DEP, which is fairly good soluble in water (1 g/L) and there is a trend in decrease of water solubility with increasing length of the side chain and molecular weight as expected. The Log Kow is lower than 3 for DEP (around 2.5), between 4-6 for DBP and BBP and values higher than 6 from DnHP onwards. Although the differences in for example water solubility and log Kow are rather high between the phthalates with a C2 backbone and those with higher C backbones, it can be concluded that there is a clear trend across the category for these two properties. For the other properties no clear trend can be seen.
The water solubility and log Kow can give an indication for the kinetic behaviour. However, the oral absorption data do only limitedly show the same trend as was seen for water solubility and log Kow. An explanation for this could be that phthalates with higher carbon

backbones (with low water solubility) will hydrolyse in the gut to a monoester, which can become more systemically available.

Evaluation of reproductive toxicity

For all the phthalates in this category there are data, which can be used to cover the endpoint reproductive toxicity. The quality of the data however differs; for the phthalates that have been evaluated in a framework there are extensive studies (2- or 3-generation studies). For the other substances several studies (repeated dose toxicity studies, continuous breeding studies, and developmental studies) have been used to come to a conclusion for reproductive toxicity. It is difficult to compare the reproductive effects observed in different studies because different protocols were used and the number and selection of endpoints investigated differ among them. Furthermore, the severity of observed toxic effects could be affected by the species and strain used, time and duration of dosing and the dose itself. For some phthalates only studies were available, in which very high dose levels were used and it was therefore not possible to identify possible reproductive effects due to severe maternal or embryotoxicity. Although it is realised, that based on these data a full comparison may not be possible, this example is still useful to get insight in the aspects to deal with for the evaluation of this endpoint within a category approach.

From Table 4.2 it can be concluded that from the phthalates selected, DEHP (side chain length C6, branched, 2 identical chains), DBP (side chain length C4, straight, two identical chains) and BBP (side chain length C4, straight, non-identical chains) cause significant reproductive effects in rodents, such as reduction of fertility of both sexes, effects on reproductive organs (particularly in males, and more pronounced if exposure occurred during from late gestation to sexual development), and teratogenic effects. The most critical effects are effects on male reproductive organs. The type of effects they produce, especially in developing male reproductive organs and the absence of clear estrogenic effects in in vivo studies (DBP and BBP exhibited weak ER-mediated estrogenic activity in some in vitro studies), leads to conclusion that all three substances act mainly by an anti-androgenic mechanism. DEHP, DBP and BBP are classified as reproductive toxicants according to EU criteria for fertility and developmental effects. However, also DnHP (side chain length C6, straight, 2 identical chains), for which only very limited information was available seems to be able to produce effects in mice (reduced fertility of both sexes and embryolethality) similar to those observed with DEHP and at comparable dose levels. Unfortunately, in the study with DnHP only high dose males were necropsied and therefore it was not possible to determine whether effects on male reproductive organs also occurred at lower dose levels.

In a one-generation study DINP (mixture of C6-C7-C8-C9 backbones) showed a statistically significant increase in the mean absolute and relative right testis, left testis and right epididymis weights and the mean relative left epididymis and seminal vesicle weights in the high-dose males; in a few subacute and/or subchronic studies, slight increases (statistically

significant) of relative testes weights were also noted at high doses. Taken as a whole, no overt toxicity was observed on reproductive organs in rats.
In mice, a very high dose (5,770 mg/kg/d) leads to a decrease in testicular weight with abnormal/immature sperm forms and uterus/ovaries atrophy in the 13-week study. In the 104-week chronic study, a NOAEL of 1,500 ppm (276 mg/kg/d) can be assumed for testicular effects, based on decrease in testicular weight (relative and absolute) observed from 742 mg/kg/d. The NOAEL for systemic toxicity in male is 1,500 ppm as well.
In the two-generation study no changes in reproductive indices were observed. From those assays, no adverse effects on fertility may be anticipated.
In a recent study DINP induced a significant level of malformations on male reproductive organs, which are indicative of an anti-androgenic mechanism, and which according to the author of the study, occurred due to the presence of some phthalates with C6-C7 ester group in ortho position in the mixture (Gray et al., 2000). In the same study, DEHP and BBP were an order of magnitude more potent than DINP, while DEP was not active at all.
Regarding fertility, it was concluded within the EU HPV program, that the effects observed in the available studies do not justify classification for DINP according to the EU classification criteria.

For DIDP (mixture of C7-C8-C9 backbone phthalates, mainly C8 backbone branched with two methyl side chains) there is no indication of organ reproductive effects in 42-44 day old (pubertal) or adult rats evidenced by histological observation in repeated dose toxicity studies and the two-generation study. In the two-generation study decrease in mean percent normal sperm was observed but of low incidence and only in P1 generation. In pups (F1, F2 and in the cross fostering satellite group) decrease in testes weight and cryptorchidism in F2 high-dose offspring were observed likely due to the low body weight since no histopathological damages were observed in adult testes. There were no changes in Reproductive Indices. From those assays no adverse effects on fertility may be anticipated.
With regard to reproductive toxicity DIDP is a developmental toxicant since a decrease in survival indices was observed consistently in both two-generation studies (Exxon Biomedical Sciences, 1997; 2000) leading to the NOAEL of 0.06% (Exxon Biomedical Sciences, 2000). As no effects were seen on fertility it was concluded within the EU HPV program, that no classification according to EU classification criteria is needed.

Among the selected phthalates there was another substance, namely DnHP (C6 backbone), for which a continuous breeding protocol study in mice was available that showed a reduction in live pups at the lowest dose level of 380 mg/kg bw/day. At this dose level no changes in body weight were observed. Only one litter was sired at the next dose level of 800 mg/kg bw/day and none at 1670 mg/kg bw/day. A crossover mating trial showed that at the highest dose levels both male and female fertility is markedly reduced. However, at this dose level systemic toxic effect, such as increase in relative liver weight and kidney/adrenal weight, was observed, but no histopathological changes. A reduction in weight of testis, epididymis and seminiferous tubules, reduction in sperm number and motility and changes in sperm

morphology were observed, at the highest dose level of 1670 mg/kg bw/day (males from other dose groups were not necropsied). Despite the limitations of this study, the effects observed are comparable to effects observed in mice after treatment with DEHP, although DEHP appears to be more potent.

For the other selected phthalates (DEP with side chain length C2, DPHP with side chain length C7 and DnOP with side chain length C8), although some effects were seen in some studies, these were observed only at high doses or they were not considered severe enough to classify them as reproductive toxicant (IPCS, 2003; NTP, 2003b; OECD, 2004a). In a repeated dose study the oral administration of DPHP to rats at dietary concentrations of 50, 250, and 1500 mg/kg bw/day for 3 months, a period sufficient to cover the complete sperm maturation, led to no effects on the relevant reproductive organs. DPHP has also recently been assigned to the category of High molecular Weight Phthalate Esters in the OECD HPV chemicals program, and is not considered to be a reproductive toxicant (OECD, 2004a).

Studies conducted with respective monoester metabolites indicate that it is likely that the monoester metabolites actually produce these effects (ECB, 2004a; ECB, 2001). However, the role of other metabolites is not fully elucidated yet. For example, in the case of DEHP its metabolites 2-ethylhexanol (2-EH) and 2-ethylhexanoic acid (2-EHA) might contribute to the teratogenic effects observed in animal studies (Ritter et al., 1987; Pennanen et al., 1993).

4.1.3.3 Applicability domain and trend analysis

In the context of a chemical category the applicability domain is a concept to identify the ranges of physico-chemical and in this case human toxicological (reproductive toxicity) properties within which reliable estimations can be made of missing data points, by the use of trend analysis, read-across, SARs and QSARs.

Physico-chemical properties:

When considering the physico-chemical properties of the selected phthalates, as has been discussed in subsection 4.1.3.2 under the chain-length approach, it appears that they do not play an important role in reproductive toxicity as it appears to be no straightforward trend and “cut-off” value, from which to derive a conclusion on why certain phthalates do and others do not cause reproductive toxicity. Also when considering the available information on oral absorption and metabolism, no conclusion can be drawn.

Reproductive toxicity

For five of the phthalates there are data from two-generation studies available. In the following figure these are indicated by grey boxes, whereas white boxes imply that such data are not available.

Figure 4.3 Overview of the availability of two-generation studies (indicated by grey boxes) and ways of filling in data gaps in (indicated by white arrows, see also Figure 2.1)

To come to an estimation of the reproductive potential of DEP, situated at the outer boundary, read-across from DBP can be applied. Looking at the physico-chemical properties of these two substances DEP is 100-fold better soluble in water and has a log Kow which is a factor 177 lower than DBP. However, there is no difference in the oral absorption in rats, which is high for both substances (90%). They are both metabolised into a monoester. The only difference in their by-products is the chain length (C2 versus C4 both linear). Considering these aspects it would be justifiable to conduct read-across from DBP to DEP. When looking at the experimental data on reproductive toxicity it was shown that DBP is a reproductive toxicant and is also labelled accordingly. For DBP there is strong evidence that it exerts its effects predominantly by an anti-androgenic activity. This activity is not mediated by binding to the androgen receptor (AR), but rather by inhibition of testosterone production (Mylchreest et al., 1999; Parks et al., 2000; Moore et al., 2001). For DEP it was concluded from the study from Gray et al. (2000) that it has no anti-androgenic activity. By applying read-across without having data on this specific activity the conclusion would be that DEP also has anti-androgenic activity. The same critical effect levels and labelling as for DBP would then have been applied to DEP.

For DnHP a prediction would be possible by applying intrapolation from both BBP and DEHP. Regarding their physico-chemical properties BBP is the most soluble in water, followed by DnHP which is a factor 56 less soluble and DEHP which is a factor 933 less soluble than BBP. Regarding the log Kow the value for BBP is about 2 log units lower than DnHP and 3 log units for DEHP. BBP and DEHP have high oral absorption percentages (about 80%), whereas no data are available for DnHP but it seems likely that the absorption value would be in the same order of magnitude.
BBP is partially hydrolysed by intestinal esterases, primarily to MBuP and benzyl alcohol with monobenzyl phthalate and n-butanol as minor products of hydrolysis. There is a preference for hydrolysis of benzyl ester, resulting in a preponderance (approximately 3:1) of monobutyl phthalate in the urine. No oral toxicokinetic data have been reported for DnHP. However, as other phthalates are converted to monoesters and alcohols and rapidly excreted, it is anticipated that DnHP would behave in the same way. N-hexanol is a metabolite of DnHP. Hexanol is oxidised to the fatty acid and metabolized by the fatty acid oxidation. The first step in the metabolism of DEHP is hydrolysis by lipases to MEHP and 2-EH (a step common in all investigated species). In principle it can be concluded that metabolism follows the same path for all three substances. From this it can be concluded that metabolism in general will follow the same path, but that different by-products can be formed. How these

will influence e.g. reproductive toxicity of these three phthalates cannot be predicted on forehand.
Based on the available data, BBP and DEHP showed to be clear reprotoxicants, which would lead to the conclusion that DnHP also has potential for this endpoint. In studies with mice DnHP produced similar effects (reduced fertility of both sexes and embryolethality) compared to those observed for DEHP. A conclusion on the potency of DnHP is difficult to draw as only high dose males were necropsied.

For DPHP a prediction would be possible by applying read-across from the mixture DINP. There are no large differences in the physico-chemical parameters. For DPHP there is no information on oral absorption and metabolism. Given the comparable physico-chemical parameters it would be justified to conclude that oral absorption is believed to be comparable for both substances. Looking at the data for comparable phthalates it seems reasonable to conclude that both substances will be metabolised to a monoester. These monoesters differ in chain length (C7 versus C6-C7-C8-C9) and branching (propyl versus n-methyl/ethyl/straight). On forehand these arguments support conducting a read-across from DEHP to DPHP, but the influence on e.g. reproductive toxicity of the different chain lengths and branching cannot be predicted. Also, as DINP is a mixture, it is difficult to interpret the experimental data as the effects seen can not be directly related to an individual phthalate with a fixed backbone. The conclusion for DINP was that it can induce a significant level of malformations on male reproductive organs, which are indicative on an anti-androgenic mechanism, probably induced by the presence of phthalates with backbone C6 in the mixture. For DPHP there are no comparable data available. A 90 days repeated dose study in rats led to no effects on the relevant reproductive organs. This study gives some indication that DPHP would not act in the same way as DINP, which provides some evidence that for the endpoint reproductive toxicity a border has to be set between phthalates with a C6 and C7 backbone.

The reproductive potential of DnOP (backbone C8 straight) could be predicted by applying read-across from the mixture DIDP (C7-C8-C9 backbone, mainly C8 branched with two methyl side chains). Looking at their physico-chemical properties they are highly uniform. Oral absorption percentages are comparable. DnOP will be converted to the monoester metabolite monooctyl phthalate and n-octanol and the mixture DIDP will also be converted to a monoester and an alcohol moiety, although for DIDP several monoesters with different carbon backbones can be formed.
Regarding their reproductive potential there are almost no data for DnOP. A continuous breeding study with mice showed a decrease of relative seminal vesicles weight at 7500 mg/kg feed. For DIDP no indications of organ reproductive effects in rats were seen. Overall it can be concluded, based on the data available, that both DnOP and the mixture DIDP do not induce the same severe effects as were seen for DEHP.

Use of alternative data (e.g. (Q)SARs and in vitro data) to define the applicability domain

Metabolism
With the METEOR program, marketed and developed by Lhasa Ltd (Leeds, UK), it is possible to make a qualitative prediction of the metabolic fate of chemical substances and it provides supporting evidence for its predictions. However, it will not give any information on the percentage that a specific biotransformation occurs. For all of the C2-C9 phthalate esters METEOR predicts comparable metabolism steps that were in agreement with data from toxicokinetic experiments. However, for example differences in absorption (e.g. phthalates) could not be predicted.

Reproductive toxicity
Maslankiewicz et al. (2005) investigated the performance of SAR-tools (DEREKfW, TSCA List of the New Chemicals Program, TOPKAT and MULTICASE) for predicting reproductive toxicity of amongst others some phthalates. Only for 1 substance (DEHP) one model (TOPKAT) did predict a structural alert for reproductive toxicity. It should be noted that this substance belongs to the training set of the TOPKAT model. For the other models no predictions on reproductive toxicity were given. For DEREKfW and the TSCA List it was clear that structural alerts for this group of phthalates were not present in the models. For TOPKAT and MULTICASE it could not be concluded what the background of the non-prediction was.

In vitro data
In Rila et al. (2005) it was concluded that for reproductive toxicity in vitro methods can be used only if a battery is developed which provides sufficient insight in the whole reproduction cycle. Such work is currently in progress in the 6th Framework Programme (ReProTect). For phthalates, which are assumed to act as endocrine disruptors mainly by interference with male hormonal system by an anti-androgenic mechanism (Mylchreest et al., 1999b; Parks et al., 2000), the time of exposure in reprotoxicity/developmental studies is the critical factor for identification of effects on developing androgen-dependent tissues and organs. The most critical exposure time in this respect was shown to be late gestation, lactation and through puberty, when the reproductive system is developing (ECB, 2004a; ECB, 2001). In developmental studies of duration 6 – 15/16 days, these effects, or at least most of these effects were not detected (ECB, 2004a; ECB, 2001). This is probably one of the reasons that, in spite of several reproductive/developmental toxicity studies on phthalates in the past 20 – 25 years, the anti-androgenic reproductive effects (such as reduced anogenital distance, hypospadias, nipple retention, malformed epididymis, et cetera) of these substances were discovered only a few years ago, when testing protocols with exposure time extended to late gestation or/and lactation were used.

Before conducting extensive reproductive toxicity studies for DEP and DPHP the application of alternative methods (SARs, in vitro methods) could be considered. However, it can be

questioned that, if these methods give no indication for a potential, the overall conclusion that they have no potential would be accepted. If the mechanism of action of DEHP and BBP would have been known, specific research on this aspect could be initiated and in first instance no full-blown studies have to be conducted.

4.1.4 Conclusions

During the evaluation of this example following the OECD chemical category concept there were some discussion points that will be elaborated more below.

Common functional groups

The C2-C9 backbone phthalate esters share the same structural features and have common functional groups (diesters with alkyl backbones differing in chain lengths/structure). As shown in the example with the phthalates the non-similarity in reproductive toxicity was mainly driven by the alkyl backbone length, which shows that not only commonality is important but also identity. When evaluating common functional groups as mentioned in the OECD guidance the length and also the composition of side chains have to be taken into account (straight, branched, aromatic). The phthalates share common functional groups (benzene carbonyl, alcohol) but differ in the alkyl chains (e.g. BBP with an aromatic ring in the side chain, DnHP straight side chains with six carbons vs. DEHP with a side chain of six carbons but branched with an ethyl group).

Also special attention has to be paid during the evaluation of chemical mixtures. In Fabjan et al. (2006) mixtures of phthalates with considerable amounts of shorter chains were discussed. It was concluded that mixtures of phthalates containing considerable amounts of C4-C6 phthalates will cause reproductive effects similar to neat phthalates. In the OECD HPVC guidance document the need for more guidance on this aspect is mentioned. The UK-HSE did mention the necessity of comparing purity and impurity profiles that can influence overall toxicity.

Likelihood of common precursors/breakdown products

It seems that there are no great differences in physico-chemical properties, absorption rates and metabolism between the phthalates investigated that could fully explain the difference in reproductive toxicity.

A major limitation of this approach is that deliverance of toxicokinetic experiments are not required in the frameworks described in this report. However, metabolic and toxicokinetic evidence is necessary for the verification of a category, especially in the absence of adequate comparable toxicity data. This problem is also identified by the OECD in their guidance.

Besides this the following remarks can be made:

This step should not only focus on the parent compound. In the case of the phthalates it was shown that possible metabolism products (e.g. BBP and the formation of benzyl alcohol, DEP

and the formation of ethanol) can be formed that have different toxicological profiles than the primary metabolites.
Furthermore, this approach cannot be used on its own to predict similarities in substance properties. The phthalates all have comparable metabolism steps leading, however, to different breakdown products. The differences in reproductive toxicity cannot be explained from the differences in these breakdown products.

Incremental and constant change across the category
Regarding their physico-chemical properties some are similar (melting/boiling point, vapour pressure) and others show a decremental (water solubility) or incremental (log Kow) trend. From the experimental analysis performed it appears that it is very likely that phthalates with the ester bound in ortho position and alkyl side chains length from C4 – C6 will produce severe reproductive effects, including effects on fertility, male reproductive organs and developmental effects other than effects on reproductive organs. Concerning the ortho phthalates with side chains length equal to or lower than C3 or equal and higher than C7, it seems less likely that they would produce severe reproductive effects similar to those observed in DBP, BBP and DEHP at low doses, but there are indications, that at very high doses these could be observed. DnHP, with very limited data on reproductive toxicity shows to some extent similar effects in rodents as DEHP.
It can be concluded that although physico-chemical properties show a clear linear trend, the endpoint reproductive toxicity does not follow a linear trend, but rather a parabolic one.

For the phthalates it became obvious that the mechanism of action has to be elucidated. In reproductive toxicity studies a wide range of effects can be found. The C4-C6 phthalates seem to cause their effect by an anti-androgenic mechanism. Effect levels have to be derived for each critical effect and needs to be addressed as such to compare other chemicals in the category for this effect. A NOAEL based on nipple retention or malformation of the epididimys may be a more specific effect than atrophy of the testis. Furthermore, for the endpoint reproductive toxicity the period of exposure is very important. The phthalates exhibit specific effects on male animals when they are exposed in the gestation period from 12-21 days.

Implications for regulatory purposes
Classification and labelling
Within this category the phthalates with backbones C4 to C6 have been classified in the EU for reproductive effects. However, in this case it seems that the classification depends on the criteria that have been set. The smaller and higher phthalates might express a similar mechanism but at higher concentrations. The application of the OECD chemical category concept for assessment of categories would not have helped in setting the border of turning point for classification and labelling. If under REACH information was provided on the non classified ones, this would have lead to under classification of these substances and, even more importantly, no risk characterisation would have been performed for these substances.

This latter is due to the fact that under REACH risk characterisations are only being made for those compounds that meet the classification and labelling criteria.

Risk assessment
Since there are strong indications that the substances within the boundaries of the category produce similar reproductive effects and at least in part share the same mechanism of action in terms of reproductive toxicity, consideration of the aggregate exposure to these substances should be taken into account. This was suggested also by DiGangi et al. (2002). However, as noted before for classification and labelling, under REACH a risk assessment will not be required for non classified substances.

4.2 Category Straight-chain aliphatic hydrocarbons

4.2.1 Introduction

Within the present context the example of straight-chain aliphatic hydrocarbons, i.e. C5-C9, is addressed. In contrast to the previous described example of phthalates, the toxicity database is very limited for this group of chemicals, especially for C7 to C9. Much of the available data is rather old (first half of the 20th century); existing evaluations rather rely on human experience than actual data. An evaluation of a category approach at the level of individual endpoints is therefore not possible for this group. However, the present example is chosen because it illustrates that one should always be aware of unexpected exceptions within a category, and not only with respect to the category members at either end of the category. In the present context it is of interest that initially Occupational Exposure Limits like the TLV (Threshold Limit Value, as set by the American Conference of Governmental Industrial Hygienists; TLVs are occupational exposure limits used by a large number of countries) for this group were based on a kind of category approach, until the specific neurotoxicity caused by n-hexane was observed in humans.

Nowadays grouping of these aliphatic alkanes may vary to a little extent. C2 to C4 are generally combined and C5 to C8 or C9; C9 may also be put in a category with C10 and higher. For the present purpose the straight-chain aliphatic hydrocarbons C5 to C9 are addressed. The exception of n-hexane within this group because of its specific neurotoxic potential is now well-known. However, for the present purpose, it is analysed how the category approach would have turned out for these compounds based on the knowledge and data available prior to the recognition of the specific neurotoxic potential of n-hexane.

4.2.2 Use of straight-chain aliphatic hydrocarbons

Straight-chain aliphatic hydrocarbons have a long history of use in industrial settings and their toxicity has long been considered to be similar. They find their use as individual chemicals (e.g. pentane as foaming agent; hexane as extracting solvent and cleaning agent) or in mixtures (petroleum hydrocarbons).

4.2.3 Category evaluation

4.2.3.1 Graphical presentation as a two-dimensional matrix of the straight-chain aliphatic hydrocarbons category

The physico-chemical properties of the straight-chain C5 to C9 alkanes are summarized in Table 4.5. It is clear that these properties follow a regular pattern with increasing chain length thereby meeting the general definition for a chemical category based on these properties. The used data sources for the substances in this category are summarized in Table 4.6.

Table 4.5 Physico-chemical properties of straight-chain aliphatic hydrocarbons

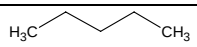
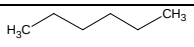
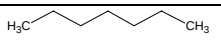
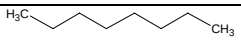
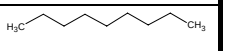
	n-Pentane 109-66-0	**n-Hexane 110-54-3**	**n-Heptane 142-82-5**	**n-Octane 111-65-9**	**n-Nonane 111-84-2**
Physico-chemical properties					
Structural formula					
Molecular weight	72.15	86.17	100.20	114.23	128.26
State of substance at 20ºC	Liquid	Liquid	Liquid	Liquid	Liquid
Melting point [ºC]	-129.7	-94.3	-90.6	-56.8	-51
Boiling point [ºC]	36.1	69	98	125.7	150.8
Relative density [kg/m^3]	626	655	688	699	718
Vapour density (air=1)	2.49	2.97		3.86	4.41
Vapour pressure [kPa]	56.6 (20˚C) 68.3 (25˚C)	16.0 (20˚C)	4.8 (20˚C) 5.3 (22.3˚C)	1.9 (25˚C)	
Surface tension [mN/m]	16.05			21.14	24.72
Water solubility [mg/l]	38	9.5-13	2.4	0.66	0.07
Log Kow	3.39	3.90	4.66	5.18	5.46

Table 4.6 Data sources for straight-chain aliphatic hydrocarbons

	n-Pentane	n-Hexane	n-Heptane	n-Octane	n-Nonane
OECD	YES	YES, CATEGORY: C6 Aliphatics Hydrocarbon Solvents - (9 CAS)	YES, CATEGORY: C7-9 Aliphatics Hydrocarbon Solvents - (9 CAS)	YES, CATEGORY: C7-9 Aliphatics Hydrocarbon Solvents - (9 CAS)	YES, CATEGORY: C9-13 Aliphatics Hydrocarbon Solvents (<=2% aromatic content) - (13 CAS)
- REPORTS		EHC HSG-IPCS UK-TR ATSDR			
- SIDS	YES	No	No	No	No
- RAR	YES	No	No	No	No
IUCLID	YES	YES	YES	No	No
HSDB	YES	YES	YES	YES	YES

4.2.3.2 The OECD chemical category concept

According to the OECD definition of a chemical category (see chapter 2 Methods and Appendix 1), the similarities should be based on the following:

Common functional groups

The straight-chain alkanes do not contain a functional group and it rather is the absence of any functional group that is the commonality. Of course two important common structural features are absence of any double bond and side chains.

Likelihood of common precursors and/or breakdown products, via physical or biological processes, which result in structurally similar chemicals.

The basic biotransformation reactions for straight-chain aliphatic hydrocarbons are rather similar although there are some differences in the rate and relative contribution of the respective pathways that are related to chain length. The longer-chain aliphatic hydrocarbons (i.e., carbon chain length > C5) are predominantly metabolized via oxidation. For n-alkanes with a carbon chain length within the range of interest, the predominant oxidation to an alcohol occurs at the penultimate carbon (omega-1 oxidation) resulting in secondary mono- or dialcohols or, through subsequent oxidation, to the corresponding ketones. Metabolites resulting from omega-2 oxidation (oxidation at C3) are found to a lesser extent. The longer-chain aliphatic hydrocarbons also may be oxidized at the C1 position, giving rise to the formation of carboxylic acids. If both a hydroxyl group and a carboxylic group are part of the same molecule lactones (cyclic esters) can be formed.

Thus in principle the basic biotransformation processes are similar for each of the C5 to C9 alkanes. The predominant metabolites are alcohols with the hydroxyl group on the omega-1 position, which may further be oxidized to the corresponding ketone.

An incremental and constant change across the category (e.g. a chain-length category, see Appendix 1 point 26).
The straight-chain aliphatic hydrocarbons form a homologous series where each member differs by a -CH2- unit. The C5 to C9 alkanes show an incremental change in several physico-chemical characteristics as shown in Table 4.6. It could also be assumed beforehand that the members of this group would have a similar toxic mode of action. The general toxic endpoints for this group of compounds are irritation (because of their defattening potential) and CNS depression (anaesthesia). The basic toxicity data are rather old and limited and do not allow an adequate quantitative comparison.

However, the available data indicate that the toxic potential (both for CNS-depression and irritation) increases with increasing chain length. This is reflected by occupational standards like the TLV (Threshold Limit Value, as set by the American Conference of Governmental Industrial Hygienists) which have been set on these effects for these chemicals. The present TLVs are based on the limited toxicological data, (industrial) human experience, and reasons of analogy. Thus it can be said that the present TLVs for n-pentane, n-heptane, n-octane, and n-nonane, are based on a kind of category approach; the present TLVs are 600 ppm, 400 ppm, 300 ppm, and 200 ppm, respectively. Initially, before its specific neurotoxic potential became acknowledged, the TLV for n-hexane was set at 500 ppm, thereby fitting in line with TLVs for the other alkanes. (The present TLV for n-hexane is one order of magnitude lower, 50 ppm.)

4.2.4 Further analysis

n-Hexane causes peripheral neuropathy also described as axonopathy; starting at the end the long nerves in the extremities and die back. This peripheral neuropathy is due to a specific metabolite of n-hexane, a gamma-diketone (2,5-hexanedione, see Figure 4.4) that cannot be formed with other straight-chain aliphatic hydrocarbons. The mechanism of action of the gamma-diketone is rather complex and cannot be predicted from its structure or from the generally applied in vitro tests. Based on the then available data on CNS-depression and irritation caused by aliphatic hydrocarbons, the toxic potential of n-hexane was considered to be in line with the others. In fact, the problems with n-hexane were recognized not before workers in the shoe industries of Japan and Italy in the 1960s and early 1970s started to suffer from peripheral neuropathy. That triggered the investigations into the mechanism of action and the neurotoxic potential of n-hexane upon repeated exposure.

It is recognized by the OECD (2005) that one should be aware of the presence of a “cut-off” point either with increasing or decreasing chain length, meaning that a category approach will only be valid within specific structural ranges. For instance, a regular changing lipid or water solubility with increasing chain length may reach a point at which the parameter exceeds a specific threshold beyond which no toxicity is to be expected, e.g. because of low reactivity or low uptake. The problem with the present group of compounds is that, despite similar

basic biotransformation pathways, the biotransformation of n-hexane leads to a metabolite with a specific toxic potential that could not have been predicted from any toxicity data obtained with one of the other category members. The basic biotransformation reactions are rather similar for all group members. The problem with n-hexane arises because omega-1 oxidation of both ends of the chain finally gives rise to the formation of the gamma-diketone 2,5-hexanedione (Figure 4.4). Gamma-diketones are known to exhibit a neurotoxic phenomenon known as “giant axonal swelling”, i.e. peripheral neuropathy. To elicit this specific neurotoxic effect the position of the ketone functions on the aliphatic chain must be 2,5-(i.e., gamma). Gamma-diketones with terminal methyl substituents (e.g., 2,5-hexanedione) and/or methyl substituents at the three and four positions exhibit the most pronounced effect. n-Pentane will be oxidized to 2-pentanol and to some extent to 2-pentanone but does not give rise to the formation of a diketone. Oxidation of n-heptane might result in 2,6-heptanone, i.e. not a gamma-diketone. Also the higher alkanes will not give rise to gamma-diketones. It can thus be concluded that toxicokinetic data of the alkanes, including n-hexane, would not have provided any suggestion for an exceptional position for n-hexane.

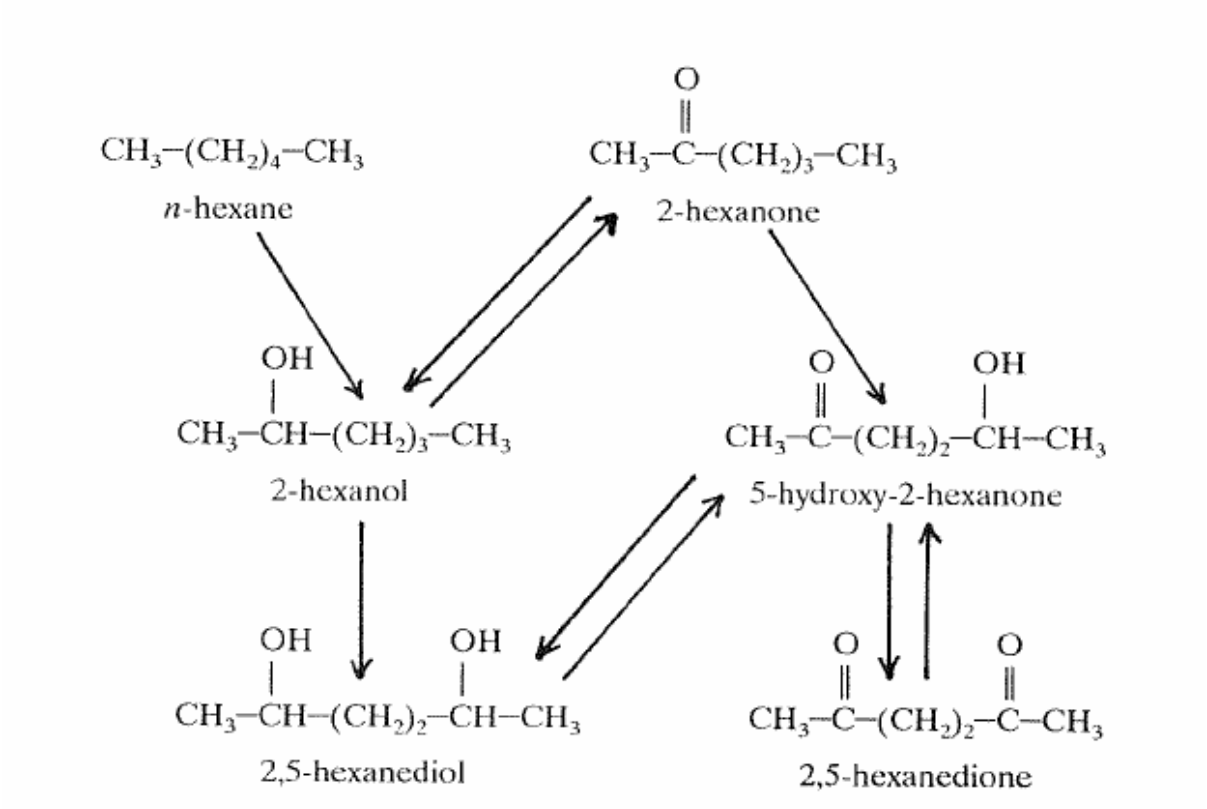


Figure 4.4 (Partial) biotransformation pathway for n-hexane

The acknowledgement of the neurotoxic potential of n-hexane also led to more detailed studies on the neurotoxic potential of the other alkanes. It appeared that n-hexane is unique in this way. Therefore, even if long-term toxicity studies would then have been available for the other members of this category it would not have provided any suggestion for the axonopathy caused by n-hexane. Neither read-across nor interpolation would have helped to recognize the

neurotoxicity caused by n-hexane. The neurotoxic potential of n-hexane only can be recognized in a repeated dose toxicity study with n-hexane itself. The gamma-diketone formed inhibits the axonal transport in the nerves, especially those in the extremities. This blockade will finally result in a dying-back of the nerves. This process will take some time and acute toxicity data or toxicity data obtained with short-term exposure (few days) to n-hexane will not reveal this effect. CNS-depression by n-hexane following acute exposure is in line with that for the other alkanes of the C5 to C9 group.

Since the axonopathy induced by n-hexane can only be detected by a repeated dose toxicity study with n-hexane itself none of the other proposals on grouping chemicals in categories (e.g. by Canada or the UK-HSE; see section 3.2) will provide sufficient guidance for the recognition of this effect. This example clearly demonstrates that the principle that “Within a series of structurally similar new substances, the two substances at either end of the series may be the only two that need to be fully tested”, as put forward by the UK-HSE, should be applied with caution.

Whether or not the peripheral neuropathy by n-hexane already can be observed in a 28-day toxicity study or that a longer exposure is required, is not clear. In most of the available studies with n-hexane animals are exposed for up to 11 to 13-weeks. Histopathological examination in rats exposed to 1500 ppm for up to 30 weeks (9-10 hours a day, 5 days a week) showed no effect on tibia1 nerve branches. Exposure at 5,000 ppm for 7 weeks was without effect. Pathological alterations (giant axonal degeneration) were seen in rats exposed to 2,500 ppm for 30 weeks or to 5,000 ppm for 14 weeks, but not in rats exposed for 7 weeks. In another study, a marked decrease in grip strength and “slowness of action” were observed in rats exposed to 3,000 ppm from exposure week 12 (12 hours a day). After week 8, motor nerve conduction velocity in the 3,000 ppm exposure group was significantly reduced. At the end of the 16-week exposure period, paranodal swellings and demyelination as well as remyelination of the peripheral nerve were observed in the 1200 ppm exposure group but not in the 500 ppm group. Motor nerve and mixed nerve conduction velocities were significantly decreased in rats exposed for 12 hours a day to 3000 ppm n-hexane by 4 weeks and became progressively slower during the study. Distal latencies (time from onset of stimulus to recording of response at the distal nerve end) were increased. Clinical signs were not observed before exposure week 10.

It appears that the induction of axonopathy depends on the exposure concentration, the total exposure duration but also on the daily exposure duration. A six-hour exposure as recommended in OECD Test guideline 412 might not be sufficiently long to detect the axonopathy within 14 to 28 days. In any case, specific attention has to be paid to measuring the nerve conduction velocity in the nerves in the extremities as well as to performing histopathological examination of these nerves.

4.2.5 Conclusions

The group of C5 to C9 straight-chain aliphatic hydrocarbons meet the OECD chemical category concept for the application of a category approach. These compounds form a homologous series where each member differs by a -CH2- unit, have structural similarities and their basic biotransformation pathways are similar. Although even basic toxicological data for these compounds are insufficient for an evaluation for every relevant endpoint it can be concluded from the available data and human experience that the category approach might be applicable for acute endpoints. The potential for inducing irritation and CNS depression is expected to increase with increasing carbon chain length and n-hexane will fit into this line. However, the difference lies in the induction of peripheral neuropathy upon repeated exposure. Long-term exposure data to either n-pentane or n-heptane also do not provide any information that might point at an exceptional position for n-hexane. Further, as has been stated before, the biotransformation of n-hexane basically does not deviate from that of either n-pentane or n-heptane. Therefore, toxicokinetic data, i.e. insight in the metabolites formed, might have provided further basis for a category approach for this group of C5-C9 hydrocarbons. In any case only a repeated dose toxicity study, including the monitoring of adequate neurological parameters, could have revealed the exceptional position of n-hexane.

Similarly, the UK-HSE strategy would not have revealed this aspect. An acute oral toxicity study and an Ames test would not have provided any indication in this direction. The principle that within a series of structurally similar substances, only the two substances at either end of a category may need to be fully tested possibly does not have a general applicability. The idea that exceptions, if present, only will be found at either end of a category should be applied with caution, as is illustrated by the present example. Although it cannot be excluded that this example is an exception in itself, it is strongly recommended not to rely on this principle by default but to be always aware of exceptions within a category.

4.3 Category Butenes

4.3.1 Introduction

An example of a category of isomeric substances and their mixtures is the butenes category. This category was evaluated and agreed upon in the OECD HPV SIAM in 2004. Before this, there was a lot of discussion between the sponsor country (the Netherlands) and the sponsor company on certain endpoints. The experiences during the evaluation of this category are described in this chapter.

The butenes category includes four isomers; two structural isomers: 1-butene and isobutylene, two geometric isomers: cis-2-butene and trans-2-butene and two mixtures. The category will be evaluated in accordance with the methods described in chapter 2. The OECD chemical category concept will be taken into account. Not all substances within the butenes category are data-rich, which makes it difficult for this example to investigate the aim of this

report. However, the surplus value of this example is that it gives a good picture of daily practice in the evaluation of categories.

4.3.2 Use of butenes

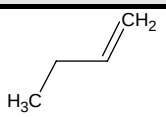
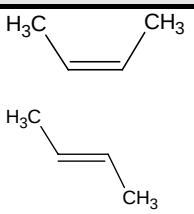
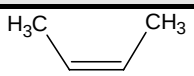
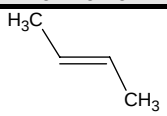
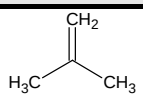
Butenes are produced commercially and are all used as intermediates.

4.3.3 Category evaluation

4.3.3.1 Graphical presentation as a two-dimensional matrix of the butenes category

The Butenes Category includes six CAS numbers and consists of substances that are similar from a process perspective. The six substances share relatively similar physico-chemical properties (Table 4.7).

Table 4.7 Physico-chemical properties of the butenes category

	1-Butene **106-98-9**	**2-Butene** **107-01-7**	**Cis-2-Butene** **590-18-1**	**Trans-2-Butene** **624-64-6**	**Isobutylene** **115-11-7**	**Butene, mixed isomers** **25167-67-3**
Structural formula		 (contains cis- (70%) and trans- forms (30%))	 (cis- form)	 (trans- form)		Contains all butene structures
Molecular weight	56.11	56.11	56.11	56.11	56.11	56.11
State of substance at 25ºC	gas	gas	gas	gas	gas	gas
Melting point [ºC]	-185.3	-138.9 to -105.5	-138.9	-105.5	-140.4	-185.3 to -105.5
Boiling point [ºC]	-6.2	0.8 to 3.7	3.7	0.8	-6.9	-6.9 to 3.7
Relative density [kg/m^3]	0.5879	0.6042 to 0.6213	0.6213	0.6042	0.588	0.5879 to 0.6042
Vapour pressure [hPa]	2,906	2,106 to 2,306	2,106	2,306	2,973	2,106 to 2,973
Water solubility [mg/l]	222	265 to 700	700	265	263	222 to 700
Log Kow	2.40	2.31 to 2.33	2.33	2.31	2.34	2.31 to 2.40

The data availability for human toxicological and environmental endpoints for these butenes is shown in Table 4.8 and Table 4.9, respectively. Most data are available for isobutylene (CAS nr. 115-11-7) and 1-butene (CAS nr. 106-98-9). For human toxicological endpoints no data are available for the mixed isomers and cis/trans-2-butene.

Table 4.8 Data availability (human toxicological endpoints) for the butenes

Test	***1-Butene*** 106-98-9	***2-Butene*** 107-01-7	***Isobutylene*** 115-11-7
Acute Oral	No data	No data	No data
Acute Inhalation	No data	LC50: male/female rat > 23.1 mg/l for 4 h	LC50: rat = 620 mg/l (270000 ppm) for 4 h LC50: mouse = 415 mg/l for 2 h
Acute Dermal	No data	No data	No data
Repeated Dose	vapour inhalation; 28 days; 6 hours/day, 7 days/week; Doses: 0, 500, 2000 and 8000 ppm; Method: OECD combined study TG422; NOAEL: = 8000 ppm	male/female rat; inhalation; 2 weeks premating, during mating and until gestational day 19; 6 hours/day; 7 days/week; Doses: 0, 2500, 5000 ppm; Method: other: OECD guideline 422 (draft 1992, final 1996) Combined repeated dose toxicity and reproductive/developmental toxicity test. NOAEL: = 2500 ppm	Isobutylene was not toxic to rats or mice exposed to concentrations up to 8,000 ppm (18,400 mg/m3) for 14 weeks or 2,000 ppm (4,600 mg/m3) for 105 weeks. For isobutylene the NOAEL is based on minimal effects on nasal cavity at the highest dose the NOAEL amounts therefore 2,000 ppm (4,600mg/m3). Other effects: increased liver and kidney weights (no dose-effect relation, no histophatological findings)
Genotoxicity (in vitro)	Ames test: negative (+/- S9, incl. *E. coli*)	Ames test: negative (+/- S9) Chromosome aberrations: negative	Mouse lymphoma assay: negative Mouse embryo fibroblast derived cell line: negative Ames test: negative
Genotoxicity (in vivo)	Micronucleus Test: negative	No data	Micronucleus assay; male mouse; inhalation 1000, 3260, 10,000 ppm in air; Result: negative
Reproductive toxicity	rat; OECD Guideline 422; other: vapor exposure; females day 1-19 of gestation; 28 days; 6 hours/day, 7 days/week; 0, 500, 2000 and 8000 ppm; Parental: = 8000 ppm	rat; other: OECD guideline 422 (draft 1992, final 1996) Combined repeated dose toxicity and reproductive/developmental toxicity test. Used in SIDS; inhalation; 46 days (male) or until day 4 of lactation; premating and mating period and until gestational day 19 (female); at total of 46 days (male); 6 hours/day; 7 days/week; 0, 2500, 5000 ppm; Parental: = 2500 ppm F1: >= 5000 ppm	No data
Developmental toxicity	OECD 422; rat (vapor) NOAEL Maternal toxicity: = 8000 ppm NOAEL Teratogenicity: > 8000 ppm	No data	OECD 414; rat (vapor exposure) NOAEL Maternal toxicity: > 8000 ppm; NOAEL Teratogenicity: > 8000 ppm

Table 4.9 Data availability (ecotoxicological endpoints) for the butenes

Matrix of available and adequate data on butenes: isomers and their mixtures						
Test	***1-Butene*** 106-98-9	***2-Butene*** 107-01-7	***Isobutylene*** 115-11-7	***Butene, mixed isomers*** 25167-67-3	***cis-2-Butene*** 590-18-1	***trans-2-Butene*** 624-64-6
Environmental Fate						
Biodegradation	RD	No data	RD	No data	RD	RD
Ecotoxicity						
Acute Fish [mg/l]	17.5 (c)	No data	19.9 (c)	No data	20.4 (c)	21.3 (c)
Acute Daphnid [mg/l]	19.3 (c)	No data	21.6 (c)	No data	22.3 (c)	22.4 (c)
Acute Alga [mg/l]	12.3 (c)	No data	13.9 (c)	No data	14.2 (c)	14.8 (c)
Terrestrial [earthworms, mg/kg soil]	259.9 (c)	No data	271.2 (c)	No data	273.1 (c)	277.0 (c)

(c): calculated with ECOSAR using the class "neutral organics"
RD: estimated to be readily degradable by using EPIWIN

4.3.3.2 The OECD chemical category concept

Primarily, the discussion with industry was focussed on identity, toxicokinetics and toxicological endpoints. Therefore, chemical category concept, as described before, can be used to reflect the discussion on these points.

Common functional groups

Each substance within this category is a C4 olefin or a mixture of selected C4 olefins that are produced from a reaction and/or separation activity in an olefins chemical plant. Four CAS numbers describe different C4 isomers; each is a hydrocarbon with the same chemical formula and one double bond between two carbon atoms. Two CAS numbers describe mixtures of C4 olefins that contain either two or all four different isomers.
The substances in this category do not, apart from the double bond, contain a functional group, which can be seen as the commonality.
The composition of the mixtures in this category was not completely clear, especially for the mixture with CAS number 107-01-7. This CAS number refers to 2-butene, which contains usually 70% cis-2-butene and 30% trans-2-butene. However, without reference differences in percentages of each individual isomer could have been present. The composition of this isomer mixture could therefore differ in the various toxicity studies.

Likelihood of common precursors and/or breakdown products, via physical or biological processes, which result in structurally similar chemicals

First, it has to be noted that in order to evaluate this likelihood, data on metabolism and toxicokinetics are needed. Typically, these types of studies are not regulatory required and therefore would require a sponsor of the chemical to do additional work beyond what is normally considered necessary.

Human health

In a paper describing uptake and metabolism following inhalation exposure to 300 ppm of C2-C8 1-alkenes (alpha olefins), Eide et al. (1995) reported that concentrations of 1-alkenes in blood and different tissues increased with increasing number of carbon atoms, while levels of haemoglobin and DNA adducts decreased with increasing number of carbon atoms. In an in vitro system, Vaz et al. (1998) demonstrated CYP2E1 epoxidation of cis- and trans-2-butene followed by hydroxylation to give 2-butene-1-ol. Epoxidation of the cis isomer was consistently faster than that of the trans isomer; however, the opposite is true for hydroxylation. Cavender (1994) indicates that branching does not affect the toxicity of C4 olefins, and that, in general, alpha olefins are more reactive and toxic than beta isomers. Isobutylene appears to be metabolized via cytochrome P450 to an epoxide, 2-methyl-1,2-epoxypropane (MEP) followed by hydrolysis via epoxide hydrolase to the diol, which is oxidized to 2-hydroxyisobutyric acid (HIBA). MEP has been identified as the primary metabolite of isobutylene in liver tissue of various species, including man (Cornet et al., 1995). In vitro studies indicate that mouse liver enzymes oxidize isobutylene to the epoxide in a reaction that is cytochrome P450-dependent and which is inducible by phenobarbital treatment (Cornet et al., 1991). Also, the formation of isobutylene epoxide was enhanced by inhibition of epoxide hydrolase or glutathione S-transferase, suggesting that the epoxide is further metabolized via these pathways.

Environment

Although the biodegradability of the butene isomers have not been evaluated with standard 28-day test guidelines, research studies designed largely to evaluate the metabolic pathways involved in the degradation of butenes have demonstrated that selected isomers can be degraded by bacteria isolated from soil and surface water samples. The results from these studies suggest that the butenes are subject to microbial degradation. However, biodegradation is unlikely to contribute to the overall degradation of butene isomers in the environment because they are gaseous and the primary environmental compartment to which they will partition is the air. Also QSAR predictions with EPIWIN show that the butenes of this category can be degraded rapidly.

An incremental and constant change across the category (e.g. a chain-length category, see Annex 1 point 26).

This approach can be evaluated in two steps: first the physico-chemical properties followed by evaluation of the toxicity endpoints.

Physico-chemical properties

The members of the Butenes Category differ minimally in physical properties (see also Table 4.7):

Melting Point (°C) ranges from -105.5 to –185.3, boiling point (°C) ranges from –6.9 to 3.7, relative Density (g/cm3 at 25°C) ranges from 0.588 to 0.6213, vapour pressure (hPa at 25°C)

ranges from 2,106 to 3,080 (determined experimentally), water solubility (mg/L at 25°C) ranges from 222 – 700 and the log Kow (at 25°C) ranges from 2.31 - 2.4.
Furthermore, physico-chemical data add to an overall understanding of the potential for distribution. Both the log Kow and water solubility favour absorption via the lung and their small molecular weight and log Kow suggest that butenes are likely to be widely distributed within the body.

Evaluation of human toxicological data (OECD HPV endpoints, see also Table 4.8)
The acute toxicity data on members of the Butenes Category is limited. Overall, these limited data suggest that the members of this category have a low order of acute toxicity.

Read-across can be applied to the endpoint repeated dose toxicity from three of the substances in this category (butene-1, isobutene, and butene mixed -1 and -2 isomers) with data to characterise the three substances without data (cis-butene-2, trans-butene-2, and butene mixed isomers). The three substances with data clearly indicate that they have a low order of subchronic toxicity. Adaptive and reversible changes in the liver indicate butenes were widely distributed within the body and metabolised. No Observed Adverse Effect Levels (NOAEL) from studies for tested members over 28-days to 2 years range from 2,000 ppm to 8,000 ppm. However, it should be noted that the NOAEL for isobutylene was based on minimal effects on nasal cavity, whereas this organ was not investigated in the 28-days study with 1-Butene. This makes a comparison for the endpoint repeated dose toxicity difficult. By read-across, the untested butene substances would also be expected to demonstrate a low order of subchronic toxicity over the extent and within the range to which the tested butenes were subjected, or to the lowest concentration of the range at a minimum.

Members of the Butenes Category that have been tested do not produce mutagenic responses either in in vitro or in vivo test systems. Butene-1, butene 1 and 2 mixed isomers, and isobutylene did not induce gene mutations in reverse mutation assays conducted in S. typhimurium and/or E. coli either in the presence or absence of metabolic activation. Butene-2 was not clastogenic to rat lymphocytes in vitro. Isobutylene tested negative in an in vitro cell transformation assay using a mouse embryo fibroblast derived cell line and in a mouse lymphoma assay both in the presence or absence of metabolic activation. In addition, neither 1-butene nor isobutylene induced micronuclei formation in mouse bone marrow cells. By read-across, these data support characterising the untested members of the Butenes Category as also having a low potential for carcinogenicity. This was further supported by the results of a two year carcinogenicity study conducted by NTP (1998). Although isobutylene produced an increase in follicular cell carcinomas of the thyroid, this effect occurred only in male rats at the highest dose, i.e., 8000 ppm. Thyroid tumours did not occur in female rats nor did they occur in male or female mice. As isobutylene is not genotoxic and as the thyroid tumours only occurred in male rats at the highest dose, i.e., 8000 ppm, the mechanism for the formation of the thyroid tumours most likely has a threshold. Overall, this data support the

conclusion that isobutylene as well as the other members of the Butenes Category have a low potential for carcinogenicity.

Inhalation reproductive/developmental toxicity studies conducted with butene-1, butene-2 mixed isomers, and isobutylene resulted in a NOAEL for each study that was the highest exposure concentration tested (5,000 to 8,000 ppm). Based on the weight of the experimental evidence and the consistent absence of observed significant toxic findings, the untested members of the Butenes Category would be expected to have a low potential for chronic, reproductive, and/or developmental toxicity and cancer.

Evaluation of ecotoxicological data (OECD HPV endpoints, see also Table 4.9)
Due to the fact that substances in the Butenes Category are gaseous at ambient temperature and pressure, and are expected to partition predominantly to the atmosphere, no aquatic toxicity testing has been conducted. Therefore, structure-activity relationship (SAR) data developed with the ECOSAR model (Cash and Nabholz, 1990) were used to address selected acute and chronic endpoints for three aquatic trophic levels (The ECOSAR model was accessed in EPIWIN (1999)). The ECOSAR model is a reliable and valid SAR model to apply to the butenes because it is based on a related chemical dataset that calculates the toxicity of neutral organic hydrocarbons whose toxic mode of action is non-polar narcosis. For this reason, no further testing was needed nor was the presentation of data from close analogues necessary. The ECOSAR model was used to predict the aquatic toxicity of butene isomers using the equation for neutral organics, a reliable estimation method for this class of substances (see further under 4.3.3.3).

Calculated log bioconcentration factors for the butenes range from 1.08 to 1.15 (EPIWIN, 1999), using log Kow values from Table 4.7. These data suggest that category members have a low potential for bioconcentration in aquatic species and are not expected to bioaccumulate.

4.3.3.3 Applicability domain
Physico-chemical properties
Read-across techniques can be applied to pure isomers within a category. For example, the log octanol-water partition coefficient (log Kow) of cis-2-butene is 2.3. If log Kow data were not available for trans-2-butene, based on the structural similarity, the data for cis-2-butene could be used to estimate the log Kow of trans-2-butene, e.g. log Kow is 2.3. Selected properties of the pure isomers can also be used to characterise a mixture containing the isomers. Using butene isomer log Kow data as an example, if only 1-butene and isobutylene data were available, 2.4 and 2.3, respectively, those data could be used to characterise the log Kow of a substance containing all the butene isomers. The log Kow value for such a substance could be represented as 2.3 to 2.4.

Human health endpoints

An example of read-across as it can be applied to human health endpoints includes the application of repeated dose toxicity studies for three of the substances in this category (1-butene, 2-butene and isobutylene) to characterise the three substances without data (cis-2-butene, trans-2-butene, and butene mixed isomers). The three substances with data clearly indicate that they have a low order of subchronic toxicity. Adaptive and reversible changes in the liver indicate butenes were widely distributed within the body and metabolised. No Observed Adverse Effect Levels (NOAEL) from studies for tested members over 28-days to 2 years range from 2,000 ppm to 8,000 ppm. By read-across, the untested butene substances would also be expected to demonstrate a low order of subchronic toxicity over the extent and within the range to which the tested butenes were subjected, or to the lowest concentration of the range at a minimum.

Environmental endpoints

The appropriateness of read-across within the butenes category would be justified based on the knowledge that the mode of toxic action for hydrocarbons is non-polar narcosis and that the toxic mechanism is disruption of biological membrane function. Therefore, these substances would exert a similar biological effect and they would be expected to do so over a relatively narrow range. Based on their similar log Kow values, an effect value (i.e., 96-hour fish LC50) for one butene isomer could be used to estimate the toxicity of another (the calculated fish toxicity values for the butene isomers range from 17 to 21 mg/L). Equally, if acute fish toxicity data were available for cis-2-butene, those data could be used to characterise the toxicity of a substance that contained cis-2-butene and trans-2-butene.

4.3.4 Conclusions

With respect to the OECD chemical category concept the following conclusions can be drawn for the butenes category:

Common functional groups

The butenes do not have particular functional groups. They are all C4 olefins with one double bond only consisting of four carbon and eight hydrogen atoms. Furthermore, they have no substituents to complicate structure or activity.

Likelihood of common precursors/breakdown products

For isobutylene and cis- and trans-2-butene experimental data demonstrated that these substances will be first metabolised by epoxidation, followed by further breakdown via hydrolase or glutathione S-transferase. Therefore it can be considered likely that the substances in the butene category will have similar breakdown products.

Incremental and constant change across the category

The physical properties of the substances in the butenes category are very similar. Their human toxicological profile is comparable. No experimental data were available for

ecotoxicological endpoints. However, (Q)SAR models were used to predict the acute aquatic toxicity of these substances, as it is considered a valid model for assessing the toxicity of neutral organics with toxic mode non-polar narcosis. From this it can be concluded that the category approach following the OECD rationales could be applied to the butenes for human toxicological and environmental endpoints. For human toxicological endpoints this was based on the low hazard profile of the butenes for the OECD HPV endpoints. For environmental endpoints this was based on the low calculated aquatic toxicity, the calculated ready biodegradability, supported with test results, low bioaccumulating potential, and the degradability in air for all butenes in the category.

With respect to the OECD chemical category concept the following remarks can be made:

- Overall it can be concluded that the substances in the butenes category meet the OECD chemical category concept for the application of a category approach.
- Alternative data, in this case (Q)SAR models, were considered valid for determining the acute aquatic toxicity, biodegradability and bioaccumulation of the substances in this category.
- For the endpoint repeated dose toxicity NOAELs ranged from 2000 to 8000 ppm, but these were based on different critical effects. The lowest NOAEL was derived from a 105 weeks study for isobutylene based on minimal effects on the nasal cavity. This organ was not investigated in the 28-days study with 1-butene. Therefore, extrapolation from isobutylene to 1-butene for the critical effect is difficult, as it cannot be excluded that 1-butene is more reactive.

4.4 Read-across for two new substances A and B and existing substance C

4.4.1 Introduction

For the New Substances framework two guidance documents on the use of (eco-)toxicological data for read-across in the notification of new chemicals were prepared by the UK Health and Safety Executive (Hanway, see Appendix 2 and 3). Probably these documents were intended for the evaluation of base-set data, although this is not completely clear from the description.

A poster was presented by Caitens and Evans, in which 4 examples of possible read-across were described. In the following one example will be discussed with respect to the (im)possibilities to use the read-across/category approach. First, in subsection 4.4.2, the example as presented by Caitens and Evans will be described. After that the methods as described in chapter 2 will be used to evaluate this example. All human toxicological endpoints, that are required for the base set level of Directive 67/548/EEC and the notified tonnage level will be evaluated. Ecotoxicological endpoints will not be considered, as the UK-HSE only presented a testing strategy for human toxicological endpoints on their poster. For this the OECD rationales will be taken into account. The needs and principles of the UK-HSE read-across proposal will also be used. The scope was to look at data-rich substances,

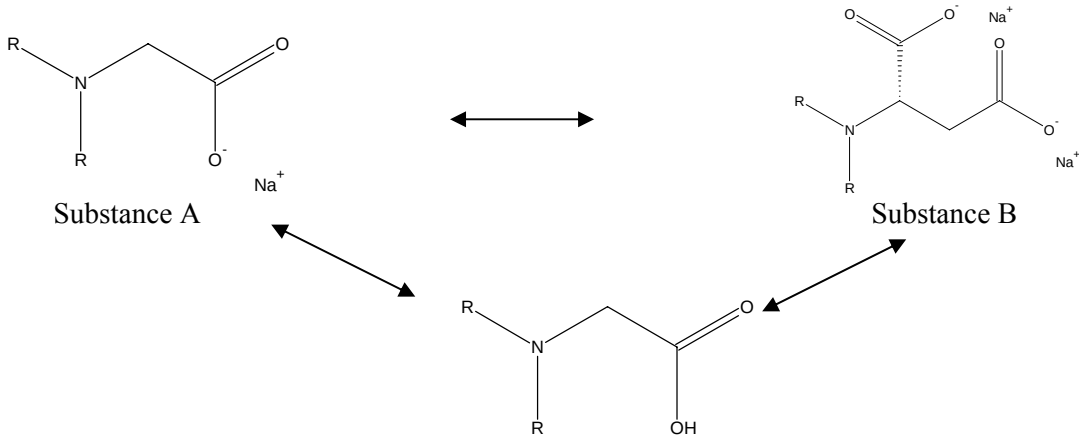
which means that first read-across between substance A and substance C will be evaluated. After that a conclusion will be made on the possible read-across with the data-poor substance B and the test plan proposed by the UK-HSE will be commented upon.

4.4.2 Description of the UK-HSE example

When high tonnages of a new substance are supplied onto the EU market, more extensive toxicological testing may be required. Such studies can be both expensive and use large numbers of animals. For example, when supply exceeds 100 tonnes per annum (tpa) the following tests can be requested:

- Fertility study
- Developmental toxicity study
- 90-day repeat dose study
- Additional mutagenicity studies

The UK-HSE presented an example, in which substance A was notified and a full base-set of toxicological information became available for this material. Substance B was then notified, some years later, and most of the toxicological data for this notification was read-across from substance A. Supply of both substance A and substance B then exceeded 100 tpa and the additional toxicological information specified above was requested to meet the requirements at this notification level. Reliable data on a third, related substance, represented as substance C was available. This is an existing substance and as such has not passed through the notification process for new substances. However, this substance was assessed within the OECD-HPVC program and a robust two-generation fertility study and a developmental study were available for this material. Taking into account the structural similarities and the data already available for these three materials, a testing strategy (Table 4.10) was set up by the UK-HSE.



R = unknown rest group

Figure 4.5 Structural formulas of the substances for which read-across was proposed

Table 4.10 Testing strategy for Substances A, B and C

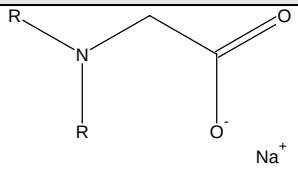
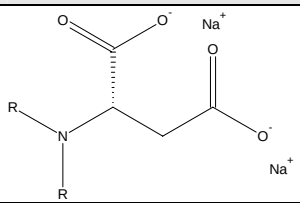
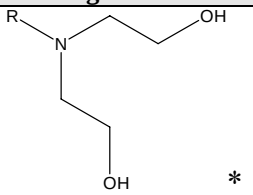
Substance A	Substance B
Full base-set testing.	Base-set read-across from substance A. Physico-chemical studies, the acute oral toxicity study and the Ames test conducted to underpin the read-across.
Read-across 90-day repeat dose toxicity study from substance B.	Conduct 90-day repeat dose toxicity study.
Read-across in vitro mammalian cell gene mutation study from substance B.	Conduct an in vitro mammalian cell gene mutation study.
Read-across 2-generation fertility study from substance C.	Read-across 2-generation fertility study from substance C.
Read-across developmental study from substance C.	Read-across developmental study from substance C.

4.4.3 Evaluation of the category or read-across approach

4.4.3.1 Graphical presentation as a two-dimensional matrix

In Table 4.11 the physico-chemical properties of two new and one existing substance are summarized. In the next table (Table 4.12) the data availability of these substances is shown for (eco)toxicological properties.

Table 4.11 Physico-chemical properties of two new substances and 1 existing substance

	New substance A	**New substance B**	**Existing substance C**
Structural formula			 *
Molecular weight	1308, (1416 (with H_2O of crystallisation))	1512.96	1165
State of substance [20ºC]	solid	solid	solid
Melting point [ºC]	> 281	> 360	> 300
Boiling point	No data (salt)	No data (salt)	No data
Relative density [kg/m^3]	1.904	1.89	
Vapour pressure [Pa]	4.6E-05	Read-across	Low in view of melting point
Surface tension [mN/m]	72	64.9	
Water solubility [mg/l]	6.2E05	>4.99E05	3.77E05
Log Kow	<-4.6	<-2.04	-2.83

* UK-HSE provided later on information that the substructure presented in chapter 4.4.2 should be corrected to the structure shown in table 4.11

Table 4.12 Data availability for human toxicological endpoints of two new substances and one existing substance

	New substance A	New substance B	Existing substance C
Acute oral toxicity [mg/kg]	>2000	>2000	>15000
Acute dermal toxicity [mg/kg]	>2000	No data	>2000
Skin irritation	Not irritating	No data	Not/slightly irritating
Eye irritation	Not irritating	No data	Slightly irritating
Skin sensitization	Not sensitising	No data	Not sensitising
Repeated dose toxicity	NOAEL 15 mg/kg bw/day (28-days study, doses 0, 15, 150 and 1000)	No data	NOAEL > 521 mg/kg bw/day (males) and > 709 mg/kg bw/day (females); 2 year feeding study with rats (doses 5.23, 52.24, 521.78 for males and 7.02, 69.33, 709.25 for females)
Ames test	Negative (+/- S9)	Negative (+/- S9)	Negative (+/- S9)
Other mutagenicity	In vitro CA human lymphocytes negative	No data	Negative in vitro CA, in vivo CA (spermatogonia), micronucleus (mouse) and dominant lethal test (mouse)
Reproduction toxicity	No data	No data	NOAEL 300 mg/kg bw/day (parental toxicity); 1000 mg/kg bw/day (reproductive toxicity and offspring toxicity), two-generation study in rats (doses 100, 300 and 1000)
Developmental toxicity	No data	No data	NOAEL 1000 mg/kg bw/day (rat, maternal and fetal toxicity, doses 100, 400 and 1000 mg/kg bw/day) NOAEL 100 mg/kg bw/day (rabbit, maternal and fetal toxicity, doses 100, 400 and 800 mg/kg bw/day)
Toxicokinetics	No data	No data	No data

For the new substance A and existing substance C full base-sets are available. Beyond that, for the existing substance C there are data on reproductive and developmental toxicity. For substance B an acute oral toxicity study and an Ames test are available for human toxicological endpoints.

4.4.3.2 The OECD chemical category concept

Common functional groups

A full evaluation of the identity of these substances is not possible since only the substructures are known. It should be noted that the reason for not presenting the full structures could be because of confidentiality issues. However, even the commonality of the substructures is not completely clear. Substances A and B consist of one or more carboxylic acids with sodium, whereas substance C consists of two alcohol groups.

Because of confidentiality no data on the purity of the substances can be given.

Likelihood of common precursors and/or breakdown products, via physical or biological processes, which result in structurally similar chemicals.
First, it has to be noted that in order to evaluate this likelihood, data on metabolism and toxicokinetics are needed. Typically, these types of studies are not regulatory required and therefore would require a sponsor of the chemical to do additional work beyond what is normally considered necessary. Therefore, the likelihood of common precursors and/or breakdown products cannot be evaluated.

An incremental and constant change across the category (e.g. a chain-length category, see Appendix 1 point 26).
The first step is the evaluation of the physico-chemical properties (in the same order or an incremental change across the category). The physico-chemical properties water solubility and log Kow of the 3 substances are compared and are in the same order of magnitude (Table 4.11). In the second step the human toxicological data were evaluated.

Evaluation of human toxicological data
For human toxicological endpoints only data on acute oral toxicity and Ames test are available for all three substances (Table 4.12). Acute oral LD50 values are above 2000 mg/kg bw for all three substances.

Substance A showed to be not irritating to skin and eyes, whereas substance C showed a slight irritation potential in some studies, but not enough to classify the substance as irritating according to EC classification criteria. Both substances were not considered to be sensitising. No data on these endpoints were available for substance B.

Substance A has a NOAEL of 15 mg/kg bw/day (28-days study, doses 0, 15, 150 and 1000), with effects mainly found in kidneys. All high dose females and one high dose male showed pale kidneys at necropsy. One female from the intermediate dose group also showed pale kidneys at terminal kill. Statistically significant increases in absolute and relative kidney weight were detected for high dose animals of either sex and intermediate dose females in comparison with controls. Many of the individual high dose female values were outside normal ranges for rats of this strain and age and the dose relationship clearly extended into the intermediate dose group. Microscopic examination of kidney sections revealed changes in high dose animals of either sex and one intermediate dose female, identified as deposits of granular eosinophilic material in the renal tubular epithelium. There were no treatment related morphological changes observed in the low dose group. For substance C a two year feeding study in rats is available with no adverse effects observed at the highest dose level (NOAEL) of 1000 ppm (i.e. circa 521 mg/kg bw/day for males and 709 mg/kg bw/day for females).

In an Ames test all three substances were considered to be not mutagenic. Substances A and C were also negative for chromosomal aberrations (no data for substance B).

For reproductive/developmental toxicity only data for substance C were available. This substance showed no evidence of reproduction toxicity or teratogenicity at the highest dose tested.

4.4.4 Conclusions

Both substances A and C are large molecules with only a few comparable physico-chemical properties, i.e. water solubility and log Kow which are also related to each other. The identity of both substances is not completely clear and there are also differences in the substructures and molecular weight. This leads to the conclusion that the similarities based on common functional groups cannot be evaluated and the needs as defined by the UK-HSE are only partly met (some physico-chemical properties are comparable).

In their needs the UK-HSE also states, that toxicokinetics need to be considered, which is essentially the same as the metabolic pathway approach as defined by OECD. No information is available and this point was also not considered by the UK-HSE in their example. Therefore, similarities based on the metabolic pathway approach cannot be judged for this category.

As concluded before substances A and C have a few comparable physico-chemical properties. The UK-HSE states in their principles that values of acute oral toxicity and Ames tests can be used to underpin the validity of read-across. When looking at the data for substance A and C it is clear that both substances show no acute oral toxicity and are negative in the Ames test. However, the major difference between the substances is the results seen in the repeated dose toxicity studies. For substance C no effects were seen up to the highest dose tested. Therefore, a LOAEL could not be derived and the NOAEL given does not give an indication of the real toxic potential of this substance (could be at least more than 1 order of magnitude higher). It can be concluded, that based on the limited results there can be a significant difference in toxic potential between substance A and C.

With respect to the OECD chemical category concept it can be concluded that a read-across between substances A and C is not justified for human toxicological endpoints, because:

- there is not enough information available to evaluate identity, purity profile and functional groups;
- there is no information available on precursors/breakdown products;
- only some physico-chemical properties (i.e. water solubility and log Kow) are comparable, which are also related to each other;
- the differences in the endpoint repeated dose toxicity cannot be interpreted;
- absence of acute oral toxicity and mutagenic potential (Ames test) do not provide additional evidence for this read-across.

4.4.5 Further analysis

Substance B has only some comparable physico-chemical properties, is different in substructure and rest group and shows identical results in the acute oral toxicity and Ames tests compared with substance A and C. In their testing strategy the UK-HSE asks for a 90 day repeated dose toxicity study and a second mutagenicity study. It has to be emphasised that when the 90 days repeated dose study is conducted for substance B it will still be difficult to interpret this study for read-across with substances A and C, as these are already showing different toxic potentials for this endpoint. Following the UK-HSE strategy, reproductive toxicity and developmental toxicity for substances A and B would then be evaluated by performing read-across with substance C. Thereby, it is implied by the UK-HSE that this could be performed based on the results of the other toxicity studies.

The category or read-across approach was applied on three substances (two new and one existing) that only have some comparable physico-chemical properties and their substructures show to some extent different functional groups. Looking at the molecular weight of the substances it cannot be concluded that the rest group for all three substances is the same. For human toxicological endpoints differences were seen in the repeated dose toxicity studies for substance A and C, which may question their similarity in mechanism of action. Regarding the proposed testing strategy, a 90 days repeated dose toxicity study with substance B will not provide enough information to judge if read-across with substance A and C is possible. Reproductive and developmental toxicity have to be evaluated on their own and not by using data from other endpoints.

5. Conclusions and recommendations

In the previous chapter examples of substances used in a category or read-across approach were evaluated using the OECD and UK-HSE guidance. This chapter describes the lessons that have been learned during the evaluation of these examples from the viewpoint of regulatory toxicology. Remarks on the suitability of alternative methods (e.g. QSARs and in vitro methods) during this evaluation will be included. Finally recommendations for incorporation into further guidance will be made.

5.1 Conclusions/recommendations

The examples described in chapter 4 were evaluated systematically using the guidance of especially the OECD. Further below the details of practical problems during the evaluation of the examples using the OECD chemical category concept are pointed out. Especially point 6 of the OECD guidance (chemical category concept), which deals with concepts on which the similarities in a chemical category can be based, will be discussed.

Common functional groups

- During the evaluation of the examples using this rationale there was much discussion amongst evaluators what is meant by “common functional groups”;
- Not only the commonality of the functional groups should be evaluated, but also the level of identity (i.e. the properties of the distinguishing groups, e.g. alkyl backbone length for phthalates);
- Within this rationale length and composition of side chains have to be taken into account (straight, branched, aromatic);
- Other aspects that need to be taken into account under this rationale are: purity profile and reactivity;
- For mixtures it is important to know how the different components relate to each other with respect to a particular endpoint (e.g. phthalates with major part C6 backbone chain (70%) vs minor part C6 backbone chain (30%));
- It is difficult to use this rationale as stand-alone to evaluate the robustness/validity of a category.

Suggestions for guidance

The first bullet of step 6 (subsection 3.2.2), dealing with common functional groups, in the OECD Manual for Investigation of HPV Chemicals should be substantiated in line with the above given comments. A definition of “common functional groups” is needed and possibly a list of common chemical functional groups could be given.

Likelihood of common precursors and/or breakdown products

- Data on absorption, distribution, metabolism and elimination (ADME) are not obligatory in most legal frameworks (New Substances, OECD HPV), making it difficult to evaluate the robustness of a category using the metabolic pathway approach;
- Focus should not only be on the parent compound or the primary metabolism product, but also on all other metabolism products, including the groups that have been split off from the main molecule (e.g. monophthalates vs benzyl alcohol);
- Formation of comparable metabolism products does not necessarily mean a comparable toxicological profile for all members of a category (e.g. biotransformation of n-hexane, n-pentane or n-heptane);
- QSARs that give a qualitative description of metabolism steps for substances within a category are only partly useful; a complete ADME model would be more helpful in relating the toxicokinetics of all category members.

Suggestions for guidance

Evaluation of the robustness of a category following the second bullet of step 6 (subsection 3.2.2) in the OECD Manual for Investigation of HPV Chemicals should include appropriate data or discussion on how the included chemicals relate to each other with respect to structure, functionality and metabolism.

Incremental and constant change across the category

- The trends for both physico-chemical and toxicological properties should be evaluated, including how they relate to each other (e.g. linear trend for alpha-olefins showing decline in acute fish toxicity with decreasing water solubility and increasing carbon chain length vs reproductive toxicity phthalates showing a linear trend in log Kow/carbon backbone length, but a parabolic trend in reproductive toxicity);
- The use of an acute oral toxicity study and an Ames test to underpin the validity of read-across (as proposed by the UK-HSE) is of limited use when none of the substances shows acute toxicity (limit test and LD50 > 2000 mg/kg bw) and all are negative in the Ames-test;
- Care should be taken during the evaluation of a particular endpoint that the critical endpoint has been investigated for all members of a category (e.g. butenes and effects on nasal cavities, anti-androgenic effects of the phthalates);
- In cases where critical effects are not caused by general toxicity the mechanism of action has to be taken into account (e.g. phthalates, straight-chain alkanes). However, this might imply the need for the conduction of specific studies that may go beyond the data requirements in many frameworks. In several cases, well conducted guideline studies will provide indications for such specific effects. This emphasises the importance of conducting high quality studies as starting point;

- The usefulness of QSARs for filling data gaps within a category approach depends largely on the validity and applicability domain of these models (e.g. QSARs for non-polar narcosis for C4 alpha-olefins vs reproductive toxicity phthalates).

Suggestions for guidance

The above mentioned points should be further discussed in guidance. Additional guidance on the use of QSARs and other models filling data gaps should be developed, particularly for human health endpoints.

5.2 Overall conclusion

During the evaluation of the examples it became clear that assessing the similarities of substances within a category approach is difficult and demands a high amount of resources with respect to time and expertise. To draw conclusions on endpoints for which no or limited data are available the need for several experts became clear (e.g. chemists, kineticists, dynamics, QSAR experts). Using all three concepts (i.e. common functional groups, metabolic pathway approach and chain-length approach) of point 6 of the OECD guidance (chemical category concept) for justifying the robustness of a category approach can increase the confidence in a category.

In order to reduce the number of animal tests/costs, read-across and category approaches should be considered before conducting tests, as is indicated in the REACH legislation. These concepts, provided they are well substantiated, have shown to work for a lot of substances that do exhibit relatively non-specific toxicological effects. However, the application of these concepts on endpoints with a very narrow time window (e.g. reproduction toxicity for the phthalates) or for substances that form different (active) metabolites (e.g. gamma-diketons) should be done with care. Also within relatively homologues series of substances differences can occur in receptor interaction in terms of potency, which in terms of REACH could have a very important impact on further handling (decisions are based on a yes/no label and/or classification class). The application of a category approach based on the metabolic approach will also be complicated by the fact that under REACH/OECD information on metabolism is not required. Furthermore, it is of utmost importance that for specific endpoints, like for example reproductive toxicity, the studies, on which the read-across/category approach are based are of high quality. However, it should be noted that each situation should be considered on a case-by-case basis and their applicability and acceptance will highly depend on how well the case is substantiated and documented. However, one should realize that even if a read-across/category approach is well substantiated, it remains a kind of inter-substance extrapolation, which inherently brings along extra uncertainties in both hazard and risk assessment. This applies with certainty to the performance of a qualitative read-across but will be even more complicated for a quantitative approach (like e.g. for NOAELs, DNELs and so on).

References

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

ATSDR (1995). Toxicological profile for Diethyl Phthalate (DEP). US Department of Health and Human Services, Atlanta Georgia.

ATSDR (1997). Toxicological profile for Di-n-octylphthalate (DnOP). US Department of Health and Human Services, Atlanta Georgia.

Cash G and Nabholz V (1990). ECOSAR. U.S. Environmental Protection Agency, Office of Pesticide, Pollution, and Toxics - Risk Assessment Division. Washington, DC, USA.

Cavender F. (1994). Patty's Industrial Hygiene and Toxicology, Vol. 2, Part B, Chapter 19, Section 2, Alkenes (olefins). Page 1240.

Cesareo D, Tosato ML, Clementi S (1987). Modelling phytotoxic activity of herbical triazines by molecular properties – a PLS approach. Environ. Tox. Chem. 6: 839-845.

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

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

CONCAWE (1998). Heavy full oils. Product dossier no. 98/109, Brussels, Belgium.

Cornet M, Sonck W, Callaerts A, et al. (1991). In vitro biotransformation of 2-methylpropene (isobutene): epoxide formation in mice liver. Arch. Toxicol. 65:263-267.

Cornet M, Callaerts A, Jorritsma U, et al. (1995). Species-dependent differences in biotransformation pathways of 2-methylpropene (isobutene). Chem. Res. Toxicol. 8:987-992.

EC (2000). Commission Regulation No 1565/2000 of 18 July 2000 laying down the measures necessary for the adoption of an evaluation programme in application of Regulation (EC) no. 2232/96. Official Journal of the European Communities 19.7.2000, L 84, 8-16.

ECB (2001). Bis(2-ethylhexyl) phthalate (DEHP) Draft Risk Assessment Report. Consolidated final report. Chapters 4-6.

ECB (2003). European Union Risk Assessment Report 1,2-benzenedicarboxylic acid, di-C8-10-branched alkyl esters, C9-rich and di-"isononyl" phthalate (DINP). 2nd Priority list. Institute for Health and Consumer Protection; Volume 35.

ECB (2003a). European Union Risk Assessment Report 1,2-Benzenedicarboxylic acid, di-C9-11-branched alkyl esters, C10-rich, and di-isodecyl phthalate (DIDP). 2nd Priority list. Institute for Health and Consumer Protection; Volume 36.

ECB (2004a). European Union Risk Assessment Report dibuthyl phthalate (DBP). 1st Priority list. Institute for Health and Consumer Protection; Volume 26.

ECB (2004b). Search Annex I to Directive 67/548/EEC on Classification and Labeling of dangerous Substances. Last updated 30/04/2004. http://ecb.jrc.it/classification-labelling/

ECB (2004c). Draft Risk Assessment Report Benzyl butyl phthalate (BBP), draft March 2004. http://ecb.jrc.it/documents/existing-chemicals/risk_assessment/draft/R318_0601_env_hh.pdf

ECPI (1997). European Council for Plasticisers and Intermediates. Information letter of D. Cadogan, to the Ministère de l'environnement and to the Institut National de Recherche et de Sécurité, on the differences between the various DINPs, Brussels, 22 sept. 1997.

Eide I, Hagemann R, Zahlsen K, et al. (1995). Uptake, distribution, and formation of hemoglobin and DNA adducts after inhalation of C2-C8 1-alkenes (olefins) in the rat. Carcinogenesis 16: 1603-1609.

EPIWIN (1999). Estimation Program Interface for Windows, version 3.04. Syracuse Research Corporation, Syracuse, NY, USA.

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

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

Exxon Biomedical Sciences (1997). Two generations reproduction toxicity study in rats with di-isodecyl phthalate (DIDP ; MRD-94-775). Project No 177535 performed for Exxon Chemical Company and Exxon Chemical Europe, 1997.

Exxon Biomedical Sciences (2000). Two generations reproduction toxicity study in rats with di-isodecyl phthalate (DIDP; MRD-94-775). Project No 177535A performed for Exxon Chemical Company and Exxon Chemical Europe, 2000.

Fabjan E et al. (2006). A category approach for reproductive effects of phthalates. Crit. Rev. Toxicol. 36:695-726.

GDCh (1992). Diethyl Phthalate (DEP). Advisory Committee on Existing Chemicals of Environmental Relevance, BUA report 104, Stuttgart.

Gray Earl Jr L, Ostby J, Furr J, Price M, Rao Veeramachaneni DN, Parks L (2000). Perinatal exposure to the phthalates DEHP, BBP, and DINP, but not DEP, DMP, or DOTP, alters sexual differentiation of the male rat. Toxicological Sciences, 58, 350-365

Hanway RH (2002a). The use of toxicological read-across data in the notification of new chemicals. Industrial Chemicals Unit, Health and Safety Executive, Magdalen House, Stanley Precinct, Bootle, Merseyside, L20 3QZ, UK

Hanway RH (2002b). The use of read-across for environmental data in the notification of new chemicals. Industrial Chemicals Unit, Health and Safety Executive, Magdalen House, Stanley Precinct, Bootle, Merseyside, L20 3QZ, UK

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

IPCS (2003). Concise International chemical Assessment Document 52 diethyl phthalate (DEP). World Health Organisation, Geneva.

Jensen AA, Nielsen PA, Lagersted A, Danielsen S and Hart J (1989). Existing Chemicals: Evaluating the use of QSAR for estimating biological effects. Nordic Council of Ministers Environmental Report 1989:9.

Maslankiewicz L, Hulzebos EM, Vermeire TG, Muller JJA, Piersma AH (2005). Can chemical structure predict reproductive toxicity? RIVM rapport 601200005, Bilthoven.

Moore RW, Rudy TA, Lin TM, Ko K, Peterson RE (2001). Abnormalities of sexual development in male rats with in utero and lactational exposure to the antiandrogenic plasticizer di(2-ethylhexyl)phthalate. Environmental Health Perspectives; 109 (3): 229-237.

Mylchreest E, Sar M, Cattley R, Foster PMD (1999). Disruption of androgen-regulated male reproductive development by di(n-butyl)phthalate during late gestation in rats is different from flutamide. Toxicology and Applied Pharmacology; 156: 81-95.

Nouwen J, Hansen B (1994). An Investigation of clustering as a tool in Quantitative Structure-Activity Relationships (QSARs). SAR and QSAR in Environmental Research. 4: 1-10.

National Toxicology Program (NTP) (1998). Toxicology and Carcinogenesis Studies of Isobutene (CAS No. 115-11-7) in F344/N Rats and B6C3F1 Mice (Inhalation Studies). Report No. NTP TR 487. National Toxicology Program, U.S. Department of Health and Human Services, Public Health Service, National Institutes of Health, Research Triangle Park, NC, USA.

NTP (2003a). NTP-CERHR Monograph on the potential human reproductive and developmental effects of di-n-hexyl phthalate (DnHP). National Toxicology Program (NTP), Center for the Evaluation of Risks to Human Reproduction (CERHR); NIH publication No.03-4498.

NTP (2003b). NTP-CERHR Monograph on the potential human reproductive and developmental effects of di-n-octyl phthalate (DnOP). National Toxicology Program (NTP), Center for the Evaluation of Risks to Human Reproduction (CERHR); NIH publication No.03-4488.

OECD (1999a). OECD Draft SIDS Initial Assessment Report for Di(isononyl) phthalate.

OECD (1999b). OECD Draft SIDS Initial Assessment Report for 1,2-Benzenedicarboxylic acid, di-C9-11-branched alkyl esters, C10-rich, and di-isodecyl phthalate.

OECD (2001). OECD Draft SIDS Initial Assessment Report for Dibutyl Phthalate.

OECD (2004a). OECD Draft SIDS Initial Assessment Report for category: High Molecular Weight Phthalate Esters.

OECD (2004b). OECD Draft SIDS Initial Assessment Report for category: Butenes.

OECD (2005). Manual for Investigation of HPV Chemicals, section 3.2 Guidance on the development and use of chemical categories in the HPV Chemicals Programme. http://www.oecd.org/dataoecd/60/47/1947509.pdf

OECD (2005a). OECD Draft SIDS Initial Assessment Report for Diethyl hexyl Phthalate.

Parks LG, Ostby JS, Lambright CR, Abbott BD, Klinefelter GR, Barlow NJ, Gray Jr. LE (2000). The plasticizer diethylhexyl phthalate induces malformations by decreasing fetal testosterone synthesis during sexual differentiation in the male rat. Toxicological Sciences; 58: 339-349.

Pennanen S, Tuovinen K, Huuskonen H, Kosma V-M, Komulainen H (1993). Effects of 2-Ethylhexanoic acid on reproduction and postnatal development in Wistar Rats. Fundamental and Applied Toxicology; 21: 204-212.

Rila JP, Bos P, Hoeven-Arentzen van PH, Maslankiewicz L and Hakkert B (2005). Applicability of (non)-validated in vitro methods for regulatory requirements. SEC report 09445A00, RIVM, Bilthoven.

Ritter EJ, Scott WJ Jr, Randall JL, Ritter JM (1987). Teratogenicity of di(2-ethylhexyl) phthalate, 2-ethylhexanol, 2-ethylhexanoic acid, and valproic acid, and potentiation by caffeine. Teratology, 35(1):41-6.

TAPIR (2005). Information Working Group 3 Non-Testing Considerations. http://ecb.jrc.it/DOCUMENTS/REACH/RIP_FINAL_REPORTS/RIP_3.3-1_INFO_REQUIREMENTS/Tapir-final_report_Appendix_9.pdf

Thomas JA, Thomas MJ. (1984) Biological effects of di-(2-ethylhexyl) phthalate and other phthalic acid esters. CRC Critical Reviews in Toxicology; 13-4, 283-317.

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

Vaz ADN, McGinnity DF, and Coon MJ (1998). Epoxidation of olefins by cytochrome P450: Evidence from site-specific mutagenesis for hydroperoxo-iron as an electrophilic oxidant. Proc. Natl. Acad. Sci. USA 95, 3555-3560.

Wold S. (1987). A strategy for ranking chemicals for biological testing. National Chemicals Inspectorate, Sweden. Rapport från Kemikalieinspektionen No. 5/87.

APPENDIX 1

MANUAL FOR INVESTIGATION OF HPV CHEMICALS (OECD, 2005)

CHAPTER 3: DATA EVALUATION

3.2 Guidance on the Development and Use of Chemical Categories in the HPV Chemicals Programme

3.2.1 Introduction

1. There are approximately 5000 chemical substances on the OECD List of High Production Volume Chemicals (last update 2004). The OECD List of HPV Chemicals serves as the overall priority list from which chemicals are selected for SIDS data gathering and testing and initial hazard assessment. The first step in making an initial assessment of an HPV Chemical is to ensure that there is adequate information on each of the elements which make up the Screening Information Data Set (SIDS). If adequate information is not available then additional data is needed to complete the SIDS for a HPV chemical.

2. For reasons of resources and animal welfare, it is important to limit the number of tests to be conducted, where this is scientifically justifiable. One approach is to consider closely related chemicals as a group, or category, rather than as individual chemicals. In the category approach, not every chemical needs to be tested for every SIDS endpoint. Rather, the overall data for that category must prove adequate to support a screening-level hazard assessment. The overall data set must allow the estimation of the hazard for the untested endpoints.

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

4. All assessments require regular review and periodic update as new information is generated. Because this is a complex area, and one in which experience is growing, the review and update of category assessments is particularly important. This will help to ensure scientifically acceptable results consistent with the original premise for the category and that methodology associated with category assessments is continually improved.

5. This document has been developed based on existing OECD SIDS cases involving categories, guidance issued under the US HPV Challenge Programme and other US EPA programmes, and on the experience gained from the OECD Workshop on the development and use of chemical categories held in January 2004. The document will be updated as further experience is gained. Furthermore, this document addresses the actual formation of categories for test plan and hazard assessment purposes. It does not address issues of presentation. These are dealt with in section 2.3.5 as well as Annex 2 (supplement 1) of Chapter 2 of this Manual.

3.2.2 Definitions and explanation of the chemical category concept

6. A chemical category is a group of chemicals whose physicochemical and toxicological properties are likely to be similar or follow a regular pattern as a result of structural similarity. These structural similarities may create a predictable pattern in any or all of the following parameters: physicochemical properties, environmental fate and environmental effects, and human health effects. The similarities may be based on the following:

- a common functional group (e.g. aldehyde, epoxide, ester, metal ion, etc.); or
- the likelihood of common precursors and/or breakdown products, via physical or biological processes, which result in structurally similar chemicals (e.g., the "metabolic pathway approach" of examining related chemicals such as acid/ester/salt); and,
- an incremental and constant change across the category (e.g. a chain-length category).

Different types of categories are described in more details in section 3.2.5.

7. The applicability domain of a chemical category identifies the physicochemical property space within which the chemical category is considered to be reliable. The applicability domain is a concept borrowed from the QSAR field[1]. In the context of a chemical category, it can be considered to identify the ranges of physicochemical, environmental, toxicological and/or ecotoxicological properties within which reliable estimations can be made of missing data points, by the use of trend analysis (interpolations and/or extrapolations), read-across, structure-activity relationships (SAR), quantitative structure-activity relationships (QSAR), activity-activity relationships (AAR) (see Annex 2 for further definitions). It can also be considered as a set of inclusion and/or exclusion rules that identify the ranges of values within which reliable estimations can be made for category members. To illustrate the concept of applicability domain, it might be observed that the category of ethylene glycols show trends in certain properties in proportion to the chain length of the glycols, but that these trends are only applicable within a *defined* range of chain lengths.

8. A chemical category can be represented graphically as a two-dimensional matrix in which different category members occupy different columns, and the different category endpoints occupy different rows (Figure 1). Data gaps can be filled in by one or more of the following procedures: qualitative read-across, quantitative read-across, use of SARs, use of QSARs[2].

9. Read-across can be regarded as using data available for some members of a category to estimate values (qualitatively or quantitatively) for category members for which no such data exist.

Qualitative read-across can be regarded as the application of SAR by using data that are internal to the chemical category. The process involves: a) the identification of a

[1] The analogy between (Q)SARs and chemical categories is made deliberately throughout this document. A chemical category can often be seen as a set of QSARs on a small scale for the different endpoints, with the advantage that all the underlying data are transparently available to the assessor. For larger categories it is possible that several different relationships can be established for a single endpoint and different members of the category (e.g. through trend-breaks) thereby defining "subcategories".

[2] together with consideration of those physico-chemical properties that determine uptake from the different environmental compartments in the case of ecotoxicity endpoints

chemical substructure that is common to two or more members of the category (which are therefore analogues); and b) the assumption that the presence (or absence) of a property/activity for a member can be inferred from the presence (or absence) of the same property/activity for an analogous member. This assumption implies that analogues behave qualitatively similarly, and is usually the result of an expert judgement evaluation rather than a more formal (mathematical) analysis.

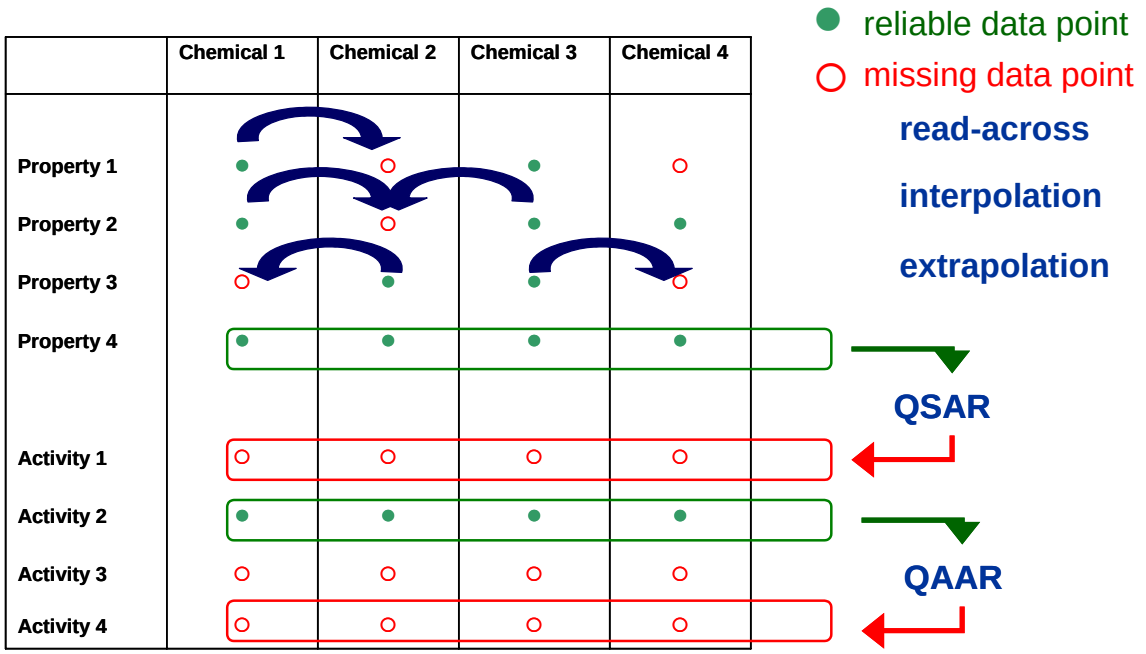
Quantitative read-across involves the identification of a chemical substructure that is common to two or more members of the category (which are therefore analogues), and the assumption that the *known* value of a property for one member can be used to estimate the *unknown* value of the same property for another member. This assumption implies that the potency of an effect shared by different analogous chemicals is similar, and is usually the result of an expert judgement evaluation as well as a more formal (mathematical) analysis.

10. Data that are external to the chemical category (data from an analogous surrogate chemical) can also be applied by using SARs. The process involves: a) the identification of a chemical substructure that is shared by a category member and by one or more surrogate chemicals; and b) the prediction of the presence or absence of an effect/activity for a category member on the basis of its similarity to the surrogate chemical. Data from surrogate chemicals should not be used selectively for only those endpoints which support the category, unless justified on a scientific basis. During the analysis to determine if the surrogate chemical is suitable, it may often be necessary to review data for endpoints other than only the endpoint of concern. However, this analysis should be performed by the Sponsor country at the initial Test Plan phase for possible review by other member countries rather than presented in the assessment documents (SIAR, SIAP). Preparing a comprehensive dossier for the surrogate chemical with Robust Study Summaries for the final SIDS documents would not be necessary.

11. QSARs can be applied by using data that are internal and/or external to the chemical category. A QSAR is a model that makes predictions of an activity (or property and in some cases the potency of the activity) from a numerical measure of chemical structure (or physicochemical property) [see also section 3.3].

12. Trend analysis can be applied when the members of a category exhibit a series of increasing or decreasing values for a given endpoint. Interpolation is the estimation of a value for a member using measured values from other members on “both sides” of that member within the defined category spectrum (see Figure 1), 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 (see Figure 1). In general, interpolation between category members is preferred to extrapolation. Especially for larger categories there may be breaks in trends, affecting the reliability of extrapolation. However, in certain cases, such as where toxicity does not change among tested category members, extrapolation to other category members may be acceptable. Interpolation can be performed with a certain confidence when the series of values is monotonic (all increasing or decreasing), but guidance and caution is needed in the case that one or more values are outliers to the trend.

Figure 1 Graphical representation of a chemical category and ways of filling in data gaps



13. Within a category different members can be selected to demonstrate the pattern or trend of interest - i.e., those selected for a category approach for environmental effects endpoints may not be suitable for assessing human health effect endpoints. Furthermore, within a category, correlations might be established for different members of the same category depending on the property (thereby establishing "sub-categories"). For example, for categories constituted of chemicals with increasing chain length, a trend might be seen for aquatic toxicity for the lower chain chemicals while a cut-off in toxicity is seen starting with a given chain length. On the other hand a correlation might be seen for another property (e.g. acute mammalian toxicity) over the whole category.

3.2.3 General Approach for Developing Categories

14. Categories accomplish the goal of the HPV Chemicals Programme - to obtain screening level hazard information - through the strategic application of testing to the category. If these test results show 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 in lieu of conducting additional screening-level testing.

15. Developing chemical categories can be considered a stepwise process (see Figure 2 for a schematic of the process and Annex 1 for examples).

- **<u>Step 1:</u> Identify proposed structure-based category and its members**

  A category can be defined in a variety of ways. Traditionally, as outlined on section 3.2.2, 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). Some categories have been defined in terms of a metabolic pathway i.e. they have a

stepwise metabolic pathway producing the different members within the category with each step.

A category definition should describe the molecular structure a chemical must have to be included in the category, including criteria such as carbon chain length, functionality, chemical or metabolite equivalence considerations, etc., and should list the specific substances covered.

The category should also be described (characterised) in terms of:

a) The relational features of the category, i.e. the chemical similarities (analogies) and trends in properties and/or activities that collectively generate an association between the members. The relational features can be regarded as the "connective tissue" that hold the category members together. Relational features include SARs, QSARs, AARs, examples of read-across, and examples of trend analysis (interpolations and extrapolations).

b) The applicability domain of the category, i.e. a set of inclusion and/or exclusion rules that identify the ranges of values within which reliable estimations can be made for category members (see paragraph 7).

Whilst the selection of a particular chemical category will normally be guided by the presence of a number of HPV chemicals in the category, it should be noted that a category may also contain other substances that are not HPV chemicals (or indeed, are not necessarily commercially available). These chemicals are legitimate candidates for the category, and may in some cases prove to be relevant candidates for further testing in order to evaluate the properties of the category as a whole[3].

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 inventories (e.g. TSCA, EINECS, ELINCS, Customs Inventories etc.) as these inventories can provide an indication as to which CAS numbers are used for marketing the substances and hence for which CAS numbers additional data might be available.

It is important that information on the purity and impurity profiles of all potential category members is collected at the same time as details of the molecular structure. Differing purity or impurities could influence the overall toxicity. For example, a category member may contain a particularly toxic impurity that is not

[3] It is recognised that in many cases the formation of a category is also dependant on which chemicals are manufactured by the consortium of companies sponsoring the category. While these considerations can legitimately influence the formation of a category, they are independent of the scientific analysis of a category and therefore not further addressed in this guidance document.

present in the other substances making it difficult or impossible to draw conclusions on the toxicity of other substances in the category. It is therefore important that category members have similar purity profiles or, where they differ, the effect of the differing purity profiles is known.

- **Step 2: Gather published and unpublished data for each category member.**
  Gather published and unpublished data on physicochemical properties, environmental fate and effects, and health effects for each member of the category. This should include all existing relevant data and not be limited to the SIDS endpoints (e.g., metabolism and cancer studies are relevant but not part of SIDS). Prepare the SIDS Dossiers for each individual category member. Specific guidance on how to prepare SIDS Dossiers for chemical categories with the IUCLID software can be found in Chapter 2, Annex 2, Supplement 1.

- **Step 3: Evaluate available data for adequacy.**
  Evaluate available data for adequacy using the OECD Guidance for Determining the Quality of Data for the SIDS Dossier (see section 3.1).

- **Step 4: Construct a matrix of data availability.**
  Construct a matrix of data availability (SIDS endpoints vs. category members) arranged in molecular weight order (or some other fashion indicating the structural progression of the category). Indicate in the cells of the matrix whether data are available or unavailable, as well as the available key study results.

- **Step 5: Perform an internal assessment of the category**

  In this step, an internal assessment of the category is performed. The internal assessment consists of:

  a) identification of the relational features that collectively generate the association between the category members. These relational features are proposed on the basis of *existing data*, which may be internal and/or external to the category.

  b) use of the relational features to fill data gaps (empty cells in the category matrix) or fill in matrix cells containing data of uncertain quality.

  In this context, the term “internal” is borrowed from the QSAR field, in which the internal assessment of a QSAR model refers to an assessment of the model performance by using the same data that were used to develop the model.

  Evaluate the category approach to determine whether there is a correlation among category members and each SIDS endpoint by looking for patterns in the matrix. The same category members do not have to be used for each evaluation, i.e., the members selected for environmental fate may be different from those used to evaluate toxicology effects.
  - If there are adequate data for a given SIDS endpoint, but no apparent pattern, the proposed category may not be appropriate and so testing may

be required for all remaining category members for that SIDS endpoint. However, an alternative category proposal may be developed e.g. the analysis might suggest that the category should be divided. (go back to Step 1).
- If there are adequate data that correlate well, the category may be appropriate and a category test plan proposal should be prepared (Step 6).
- If adequate data do not exist, but the structure-based category is reliable for one or more SIDS endpoints, then a category approach may still be proposed (go to Step 6).

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.

- **Step 6: Prepare category test plan.**

Category test plans (Step 6 of Figure) should include a category definition, rationale, and matrix of data availability (see example category test plans in Annex 1.) and be accompanied by SIDS Dossiers for each category member.

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

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

The matrix of data is an essential part of the test plan and provides a useful tool for consideration and presentation of the available data (see Annex 1). Assuming the SIDS endpoints are rows in the matrix, each row must have data in at least one cell. 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 members in the middle of the range of toxicity. Any change in a tendency for a property should be accompanied by data in the adjacent cells in order to define the limits for the resulting subsets of the category or sub-categories. Assuming the columns are the category members, one or more columns may have all empty cells, i.e. no test data available. There are no rules for the number of columns and cells that must be filled nor the number that can be empty. Acceptability of the matrix will depend on the number of members in the category, the SIDS endpoint, and the confidence in the interpolation and extrapolation.

When selecting a sample to test, it should be representative of the substance marketed, including the presence of any manufacturing impurities (see section 2.3.3).

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

At this point in the process, the sponsor country may consider to submit the test plan to the other OECD member countries for consultation.

- **Step 7: Conduct the necessary testing.**

- **Step 8: Perform an external assessment of the category and fill data gaps**

  In this step, some or all of the relational features are assessed by checking whether the predictions they make for data gaps (or data points of dubious quality) are accurate on the basis of *newly-generated* experimental data, obtained in Step 7.

  In this context, the term "external" is being borrowed from the QSAR field, in which the external assessment of a model refers to an assessment of the model performance by using independent data different from that used to develop the model.

  Add the new data to the SIDS Dossier for the relevant category member and evaluate whether the existing data and the new data support the proposed category.

  - If the results support the category, the testing phase is complete. A SIAR for the category of chemicals should then be prepared including a category analysis. The category analysis will include a summary of the one or more SIDS endpoints in which the category "holds", including the interpolation/extrapolation of test results to the remaining, untested matrix cells (see below). The SIAR would receive Member country review at the SIAM meeting.
  - If the results do not support the category return to Step 5. Further testing may be carried out, members of the category may be changed (e.g. dividing the category as appropriate), or the category proposal may be dropped altogether. The latter implies that testing will then be done to fill all appropriate SIDS endpoints for each HPV category member.

The initial assessment of the category members can only be used by member countries for purposes of classification and labelling or risk assessment if all the

data gaps are filled. As indicated in section 3.2.2, data gaps are filled by read-across, extrapolation or interpolation. This is specific to each category. No definitive guidance can be provided for the moment. A few examples are provided in Annex 1.
Available options for filling data gaps include:

1. Qualitative: it is concluded that all the members of the (sub)category do or do not possess a particular property e.g. *in-vitro* mutagenicity.
2. Quantitative: it is concluded that all the members of the (sub)category possess a particular property with a similar potency or evolving according to a regular pattern. Data gaps can be filled e.g.
    - by using the value from the closest analogue in the (sub)category;
    - by using a worst case approach i.e. using the value from the most hazardous substance in the (sub)category, or in case of interpolation, the value from the most hazardous of the two closest analogues (see figure 1);
    - by estimating quantitatively the potency of the property from the potency of the two closest analogues or from the regular evolution of this potency over the different (sub)category members.

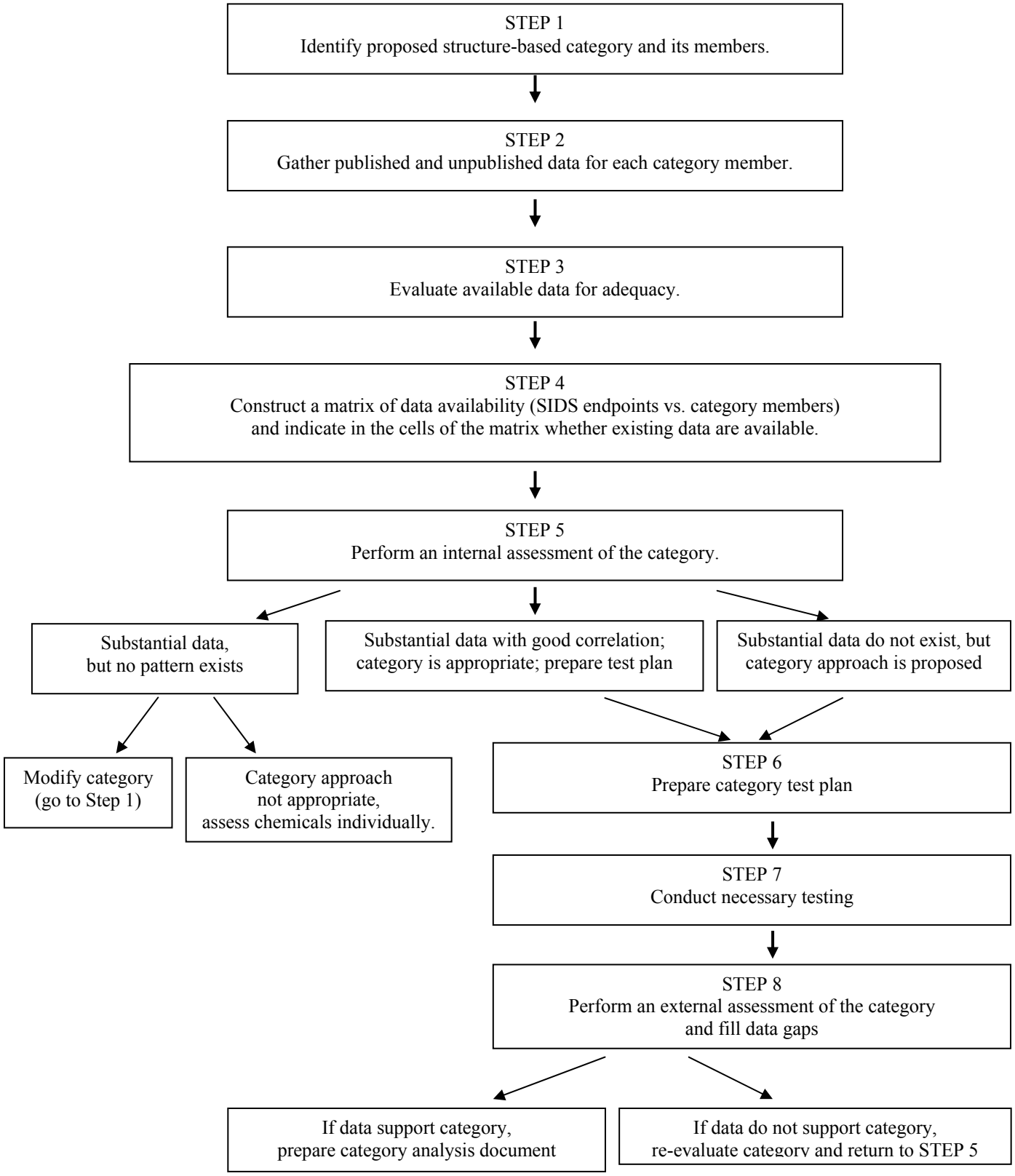
    There is currently only limited experience with quantitative data gap filling for toxicological endpoints. It should be applied with caution and the guidance will be revised as soon as more experience is available.

QSARs could be used to support proposals for filling data gaps by any of the mechanisms described above.

For categories composed of complex substances, approaches like the toxic equivalency factors or toxic units approach could be investigated to fill data gaps.

The mechanism by which the data gaps are filled in a given category should be any case be described transparently in the SIAR.

FIGURE 2: PROPOSED PROCESS FOR DEVELOPING CHEMICAL CATEGORIES



3.2.4 Use of QSARs for the Development of a Category

16. Greater confidence and further demonstration of the category approach may be gained through applying QSAR models on all category members for a given endpoint in case reliable QSAR models are available for the category members and the endpoint. QSARs can contribute at all stages of category development and consideration. Based on experienced assessment of the quality of output taking into account limitations and strengths of a range of models, QSARs may contribute not only for endpoints and compounds within categories where there are no relevant data but also in the interpretation of weight of evidence for mixed datasets and analysis of trends.

17. The output of QSAR modelling is particularly valuable in hypothesis generation and testing for step 1, "identifying the structure-based category and its members". Analysis of physico-chemical and (eco)toxicological data from the members of a category should demonstrate clear relationships between those members; outliers can then be investigated for their 'eligibility' for membership. It also permits hypothesis testing of several possible combinations and permutations for category definition. The more transparent, evolving analytical models provide access to detailed description of relevant data in the training sets. This can facilitate initial consideration of trends to establish the nature and bounds of the category.

18. QSAR modelling can also assist at this stage in defining the appropriate bounds of the proposed category, through consideration of measures of similarity for chemical descriptors in the models. In the more transparent, evolving analytical models, the bounds of this similarity can be specified and the category defined accordingly. In some cases, this can lead to the definition of a more extensive category than was originally envisaged.

19. In addition, QSAR can in some cases be used to assess similarities in metabolic pathways across the group, and this information can be helpful in assessing similarities and differences within the category.

20. Results of QSAR modelling are also relevant to step 2 ("Gathering published and unpublished data for each category member"). In addition to contributing to trends analysis for potential members of the category where no data have been identified, considered output of a battery of models can also add weight of evidence to increase confidence in trends analysis, where the pattern is not clear or consistent based on available data. For example, evaluated QSAR output may contribute where dose spacing or non-comparability of experimental protocols in available studies for different members of the category precludes meaningful analysis of quantitative trends of effect levels. In compiling this information, however, it is important to distinguish where the models contribute additionally to identified experimental data – i.e., that they are not simply duplicating the information, based on replication of its inclusion in their training set. The ease with which this information can be accessed for various models (if at all) varies, depending upon degree of transparency.

21. In relation to step 3 ("Evaluate available data for adequacy"), for QSAR modelling, this requires consideration of aspects related to the training sets and the models, themselves. Relevant aspects include criteria for inclusion of and nature of data in the training sets, the nature of the analysis for consideration of similarity, the criteria for weight of evidence for delineation of a positive/negative response and the nature of validation of the models and

aspects thereof, including concordance, sensitivity and specificity for specific endpoints and subsets of chemicals. For characterization of hazard for related endpoints, critically evaluated QSAR output can be combined with weighting of the endpoints themselves (e.g., *in vivo* versus *in vitro* genotoxicity) as a basis for meaningful contribution to hazard characterization, particularly where data are lacking or mixed.

22. For step 4 ("Construct a matrix of data availability"), then, it will be important that results of QSAR modelling be clearly distinguished from those which are based on data. As indicated above, only evaluated results of QSAR modelling which contribute additionally to weight of evidence determinations or quantitative trends analysis should be included. This would include, then, only results for modelling, where evaluated output meaningfully contributes to weight of evidence or trend analysis (this could be for substances where there are no data or where datasets for category definition are uninformative or mixed).

23. For step 5 ("Perform an internal assessment of the category"), the output of QSAR modelling introduced and considered as outlined above can contribute to trend analysis for compounds in the series both for those for which there are data and those for which there are not. Through measures of similarity, it can also contribute to delineation of the bounds of the category.

24. For step 6 ("Prepare category test plan"), where critically evaluated output of QSAR contributes meaningfully to trend analysis, it may obviate the need for testing of certain members of the category. Rationales need be based on well documented critical evaluation of the output of batteries of models, with clear delineation of strengths and limitations and take into account availability for other members of the category and consistency overall of critically evaluated QSAR output and data.

25. For step 8 ("Perform an external assessment of the category and fill data gaps"), the principles outlined above for consideration of QSAR in development of the test plan are also relevant in considering their contribution to the initial assessment. This contribution must necessarily be based on critical evaluation of the output of a suite of models, based on an understanding of their relative limitations and strengths for the specified application.

3.2.5 Guidance on different types of categories

Chain length

26. These are defined as categories showing an incremental, and usually constant, increase in chain length across the category. There is an assumption that each category member exhibits the same toxic mode of action. Examples are the homologous series of alpha-olefins (see Example A in Annex 1) where each category member differs by a $–CH_2-$ unit and the ethylene glycols where there is an incremental increase in the number of CH_2CH_2O groups.

27. Categories defined by chain length generally show an incremental change in molecular weight and other physico-chemical properties such as water solubility or Log Kow. However, not all properties will necessarily exhibit a linear relationship with chain length and care must be taken in making assumptions about such trends. For example, the alpha-olefins show declining acute toxicity to fish with increasing chain length and decreasing water

solubility. There is an apparent 'cut-off' point between the C8 and C10 chain length at which acute toxicity to fish is no longer observed due to the decreasing water solubility. For aquatic toxicity, the interplay between decreasing water solubility and increasing log Kow – a key indicator of uptake from water - with increasing carbon chain length is often important in determining this cut-off point. Similarly, a trend of increasing molecular weight may lead to decreasing systemic toxicity as absorption decreases and there may be a change of physical state of the category members as chain length increases.

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

29. QSARs can be used to help justify the category and fill data gaps. In general, substances at either end of a chain length category should have all SIDS endpoints fulfilled, preferably with test data. This permits interpolation of data to the other category members rather than extrapolation and increases confidence in the read-across. For example, a linear regression has been used to predict acute aquatic toxicity of long chain alcohols. For categories where there is more than one variable, such as variation in chain length and degree of branching of the chains, more category members are likely to be required to bring confidence to the interpolations being made.

Metabolic pathways

30. The underlying hypothesis for a metabolic series is a sequential metabolism of a parent chemical to downstream blood metabolites that are chemicals of interest. Hazard identification studies with the parent compound could then be used to identify the hazards associated with systemic blood levels of the downstream primary and secondary metabolites and once quantified, can be used in place of studies using direct exposure to primary and secondary metabolites themselves. In certain instances, the metabolism of the parent compound within barrier tissue (e.g. lung or gut tissue) occurs so rapidly that the initial primary metabolite is the predominant chemical found within the blood. Under these circumstances data from hazard identification studies conducted with that primary metabolite itself can be used to identify hazards for the parent compound. PBPK or PBPD models may help to define categories. The metabolic pathway approach is usually reserved to some toxicological endpoints. For physico-chemical properties, environmental fate and ecotoxicity, information on the parent compound would need to be available.

31. The first technical issues faced when forming a metabolic series is to determine if the metabolism that is assumed to occur does occur. This is necessary before moving any further in developing a metabolic category and preferentially should be determined *in vivo*. In certain instances, *in vitro* metabolic studies can be used to help identify metabolic pathways, but the definitive evidence should be conducted in whole animals. The primary and secondary metabolites should be detected either in the blood or tissue. Primary and secondary metabolites that cannot be readily determined in blood or tissue should not be

candidates for a metabolic series approach without some limitation placed upon the use of the information.

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

33. The third technical issue provides a limitation for the metabolic approach to forming categories. The metabolic category reasoning is only useful for identifying hazards related to systemic blood levels of the parent compound and/or primary and secondary metabolites. Other endpoints of hazard identification studies that are dependent upon site of contact effects (e.g. eye, skin, respiratory tract irritation, irritation to gastric mucosa) cannot be addressed using the metabolic category logic. These sites of contact effects are often due to the physical chemical property 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 sensitization) or are dependant upon physical chemical properties (e.g. volatility and LC50 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.

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

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

36. The following specific issues should be taken into account when developing a metabolic pathway category, according to the stepwise procedure described in section 3.2.3

- *ad* step 1: Provide definitive information on the metabolism of the parent chemical to the primary and secondary metabolite. This information should also

include, preferably, a time course data for either blood or tissue for both the parent chemical as well as the primary and secondary metabolites.

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

- *ad* 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 is not available for the metabolic series in question and a study is to be designed and conducted, then the parent compound should be tested, so that blood levels of all category members will be present. The toxicokinetic and metabolic experiments that provide the basis for the metabolic category should have robust summaries prepared and be included in the SIDS Dossier for the parent chemical, primary and secondary metabolites. Within these robust summaries a table should be included detailing the relative blood levels of the parent chemical, primary and secondary metabolites.

- *ad* step 5: A quantitative analysis between exposures of the parent chemical and the primary and secondary metabolite is not necessary as the point of the OECD HPV Chemicals Programme is to provide hazard identification studies for these materials, not a quantitative analysis as would be done for risk assessment purposes. If the chemical becomes a chemical of concern, then additional toxicokinetic analysis (including preparing a model) may be appropriate, but for the purposes of the screening level OECD HPV Chemicals Programme it is not necessary.

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

38. On the other hand the same concept underlying the metabolic pathways can be used for environmental degradation processes. For example, for a substance which hydrolyses very rapidly in aquatic test systems (half-life < 1 hour), the aquatic toxicity endpoints can be covered by the test results with the degradation product(s) [see also the OECD Guidance Document on Aquatic Toxicity Testing of Difficult Substances and Mixtures No. 23, ENV/JM/MONO(2000)6].

Chemical mixtures

39. Categories can sometimes apply to series of chemical reaction products or chemical mixtures[4] that are, again, related in some regular fashion. Analogous to the basic "discrete chemical" category model, in a mixture category some, but not all, of the individual mixtures may undergo testing. Annex 1 illustrates this using a category made up of linear alkylbenzene mixtures. This example was used to assess these chemicals in the OECD HPV Chemicals Programme. This is a relatively simple example of the type of approach that can work in the HPV Chemicals Programme. Additional guidance for other types of mixture categories is given below.

Isomers and their mixtures

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

41. Structural isomers are molecules with differences in the arrangement of their atoms, such as butene-1 and isobutene. Structural isomers can include:

- chain isomers, for example hydrocarbon chains with identical or variable lengths and variable branching patterns

- position isomers, for example hydrocarbon chains with a functional group that varies in position along the chain

42. A third type of structural isomer is referred to as a functional group isomer. These isomers also have identical molecular formula, but contain different functional groups. Examples of two functional group isomers with $C_4H_{10}O$ as a molecular formula are 1-butanal and 2-butanone. Each of these isomers contain a carbonyl group (C=O), but are representative of two different chemical families, aldehydes and ketones. Although structural isomers, this type is less likely to be considered within a category for the Programme because functional isomers can have very different chemical and biological properties. Functional isomers are not included within the scope of this guidance.

43. Geometric, or stereo, isomers contain their molecules in the same arrangement, but a section or sections of each have different spatial arrangements. For example, *cis*-butene-2 and *trans*-butene-2 each have carbon groups on either side of a double bond, which cannot rotate, that are arranged on either the same side of the molecule (*cis*-) or opposite sides of the molecule (*trans*-).

44. Geometric and select structural isomers can have similar or somewhat different chemical properties and although they can behave identically in many chemical reactions, there may be enzyme specificity towards one over another in biological systems. An example of such specificity is select carbohydrates, which may be metabolised or not depending on the orientation of functional groups.

45. There are general rules for using read-across techniques as they apply to isomers:

[4] The concept applies to reaction products or process streams and does not refer to intentional mixtures of substances (or preparations).

- Relatedness - The substance(s) without data as well as the substance(s) with data are similar such that their physicochemical, biological, and toxicological properties would be expected to behave in a predictably similar manner or logically progress across a defined range.

- Structural Similarity - The substance(s) without data possesses a small incremental structural difference from the reference substance(s) or the difference between the two would not be expected to affect the property sufficiently such that it could not be accurately predicted.

46. There can be instances within a category of isomers, specifically as related to structural isomers, when read-across for an endpoint is not appropriate. An example is illustrated with two categories of isomers other than the butenes, the pentanes and hexanes. Though the pentanes may be broadly described as isomers, they actually represent three types of hydrocarbons, normal alkanes, branched alkanes, and cyclic alkanes. It is known that n-pentane, 2-methylbutane, 2,2-dimethylpentane, and cyclopentane exhibit distinct differences in potential biodegradability. n-Pentane and 2-methylbutane are readily biodegradable, whereas 2,2-dimethylpentane and cyclopentane are poorly biodegraded. Therefore, it would not have been possible to assess the biodegradability of the poorly biodegradable pentanes if they had no data using the results from the readily biodegradable pentanes even though the pentane isomers could still be considered a category for all other endpoints within the Programme. 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 n-hexane exposure has not been demonstrated to occur with exposure to other hexane isomers and a discussion of this effect within a hexane isomer category would have to isolate n-hexane from the other isomers.

47. An example of a category of isomers is provided in Annex 1 (Example D: Butenes and their mixtures). Based on this example, general principles of read-across/extrapolation and application of data within a category of isomers and their mixtures can include:

- Select 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.

Complex substances

48. Complex substances include a diverse range of materials which are frequently described as substances of *Unknown or Variable composition, Complex reaction products or*

Biological material (UVCB Substances). There are many different types of complex substances, though generally they all have the following characteristics in common.

- They contain numerous chemicals (typically closely related isomers), and cannot be represented by a simple chemical structure or defined by a specific molecular formula. They are, however, assigned unique Chemical Abstract (CAS) numbers (see note[5] below about unique issues with CAS numbers for UVCB substances).

- They are not intentional mixtures of chemicals.

- Many are of natural origin (e.g., crude oil, plant extracts) and cannot be separated into their constituent chemical species.

- The concept of "impurities" typically does not apply to complex substances.

49. Category approaches for complex substances may vary, though generally the approach will be related to how the substances are manufactured, defined and used. For example, petroleum substances are generally defined by hydrocarbon chemistry (e.g., aliphatic hydrocarbons, aromatic hydrocarbons, etc.), physicochemical properties such as boiling range or carbon-number range, manufacturing and processing conditions, and common use categories. For hydrocarbon solvents [see example E provided in Annex 1], the categories are based on the typical chemistry and carbon-number range of hydrocarbon solvents and common uses. Under this approach, those hydrocarbon solvent substances with similar chemistry and carbon-number range are grouped together in the same category and the category is defined by the composition of those substances. This approach is practical and has the benefit of making sure that similar commercial products are grouped together in the same category.

50. Based on the example described in Annex 1, some general guidance can be provided for developing chemical categories with complex mixtures:

- It is important to clearly characterise mixtures, details of the production process can be useful. It is necessary to identify the following attributes of a complex mixture:
    - Composition (what is present and in what proportion)
    - Impurities (substances present that are not wanted but need to be identified)
- Properties of the components of a complex mixture can be applied to the complex mixture if the properties of the single components are similar.
    - It is necessary to identify representative components of the mixture to cover the carbon range and structures of the mixture.
    - Components with outlying properties need to be identified (e.g. specific toxicity of hexane compared to other aliphatic hydrocarbons, higher water solubility of aromatic hydrocarbons compared to aliphatic hydrocarbons).

[5] CAS numbers are important for identifying substances; however, for UVCB substances they do not represent a unique chemical and the specificity of the CAS number definition may vary (some CAS number definitions are rather narrow, some are very broad). CAS numbers for complex petroleum substances are based on a hierarchy of considerations including hydrocarbon type, carbon number range, distillation range and the last processing step. Because of these numerous considerations, similar products sometimes have different CAS numbers. There are also historical and geographical reasons why similar substances may have been assigned different CAS numbers. Further, some CAS numbers have a broad definition that may fit different substances that would fall into different categories. Because of this, physical properties and chemical structure are the preferred way to construct categories of complex substances.

- Properties of a complex mixture can be read-across to another complex mixture if the composition of the two are similar.
- Quantitative read-across is more difficult (ranges can be used where applicable).It is necessary to carefully consider the dose for read-across because of the nature of the mixtures and the amount of components of concern.
- It is necessary to carefully identify representative substances for testing purposes.

Metal and metal compounds

51. The concept of chemical categories has traditionally been widely used for inorganic substances. However, not much experience is available to date of a systematic use of this approach. The concept is being used for the assessment of nickel and nickel compounds (see example in Annex 1).

52. There are a number of assumptions underlying any grouping of metal compounds for estimating their biological properties. The main assumption is that it is the metal ion that is responsible for the effects to be assessed. This is considered to be a reasonable assumption for the majority of the inorganic and some organic anions. This implies that in the case of inorganic salts, the toxicity of the counter ion is assumed to be largely irrelevant in producing the effects to be assessed. If the counter ion influences significantly the effects of the compound to be assessed, it can not be part of the category. Where a metal can have different valence states (e.g. chromium), the toxicities of the different valence states may vary, and the different valence states considered separately.

53. The water solubility of the metal compounds is often used as the starting point for establishing a category, as this reflects the availability of the metal ion in the different compartments of interest. For inorganic nickel compounds, a number of sub-groups have been suggested, reflecting different ranges of aqueous solubility. In contrast to inorganic nickel compounds it is not obvious how to group organic nickel compounds based on solubilities alone.

54. Based on the example of nickel and nickel compounds, some tentative general guidance for metal and metal compounds can be proposed:

- The main assumption is that the metal ion (or ion complex) is responsible for the effects to be assessed (the toxicity of the counter-ion is assumed to be largely irrelevant in producing the effects to be assessed).
- One basis of grouping could therefore be water solubility (inorganic metal compounds), taking into account:
    - transformation/ dissolution of insoluble compounds
    - bioavailability of the metal ion in the environment
    - solubility in biological fluids
    - persistence in the body
- The assumption that the metal ion (or ion complex) is mainly responsible for the effects rather than the counter-ion may not work for local mammalian toxic effects.
- Possible differences in the toxicity of different oxidation states of the metal ion (or ion complex) should be considered.
- Whilst the assumptions shown above can be expected to be valid for a wide range of inorganic compounds, these do not necessarily apply to organically based metal

compounds. A different approach may be needed for grouping organic metal compounds. .

55. It should be noted that whilst this example considers groups of metallic cations, similar considerations would also apply to salts of anions where there are concerns for toxicity (e.g. cyanides, oxalates).

3.2.6 Experience in Developing Chemical Categories

56. OECD experience provides a framework for handling categories. However, since that experience is limited, lessons learned in the OECD, and other similar programmes such as the US HPV Challenge Programme, will provide a measure of feedback and review.

57. The largest categories applied in the OECD HPV Chemicals Programme to date contain eight to ten chemicals. This is not a formal maximum, but acceptable categories will tend to be self-limiting because endpoint trends are generally disturbed as structural variations become more complex. Practically the analysis of large categories can also become unwieldy tending to limit the size of categories proposed. In this regard, groups of related individual categories may be considered, each one contributing elements in the design and implementation of an overall category strategy. A larger category may be justifiable in certain cases, such as when toxicity of the category is generally low.

58. Annex 1 contains a number of examples of how a category approach has been used for the purpose of collecting, reporting, and assessing hazard information in the OECD HPV Chemicals Programme. Other examples of categorising chemicals for hazard assessment purposes include the CONCAWE (the European oil company organisation for environment, health and safety) approach of categorising chemicals in petroleum streams (CONCAWE, 1998), approaches to assess the ecotoxicity (Bowmer et al., 1998) and health effects (Clary, et al., 1998) of lactate esters, and a number of category/SAR analyses by the German authorities (Greim, et al., 1994, 1995, 1998; and Poelloth and Mangelsdorf, 1997).

APPENDIX 2

The use of human toxicological read-across data in the notification of new chemicals (Hanway, 2002a)

Dr Ros Hanway, Industrial Chemicals Unit, Health and Safety Executive, Magdalen House, Stanley Precinct, Bootle, Merseyside, L20 3QZ, UK.

The needs and principles that require consideration are outlined below.

3.1. Needs

3.1.1. Comparison of the purity and impurity profiles.

The similarity of the purity and impurity profiles of the new substance and the structural analogue needs to be assessed. There is a need to ensure that there are no differences in the purity or impurities that are likely to influence the overall toxicity.

3.1.2. Evaluation of physicochemical properties.

Comparison of the physicochemical properties of the new substance and its related analogue, particularly the physical form, molecular weight, water solubility, partition coefficient and vapour pressure, provides useful information as to their similarity.

Experience has shown that differences in physicochemical properties can produce marked differences in toxicity. One read-across had a 100-fold difference in the water solubility between the two substances. In this case there was also a classification change from not classified in the analogue to acutely harmful in the new substance after testing.

3.1.3. Consideration of the toxicokinetics.

The likely toxicokinetics of the substances, including the possibility of different metabolic pathways coming into play, needs to be considered.

3.1.4. The significance of reading-across results obtained from outdated test methods.

Experience of using Annex V test methods/OECD test guidelines can lead to certain revisions to their protocols. Examples of toxicology tests that have been recently revised are the skin sensitisation and 28-day oral repeat dose toxicity studies. If read-across data have not been produced using the most current Annex V test methods, particularly careful consideration of the quality and suitability of a method is important.

3.2. Principles

3.2.1. The value of the acute oral toxicity and Ames tests as a means of underpinning the validity of the read-across.

As well as providing an indication of any potential mutagenic concerns of a new substance, the Ames test is also a quick and cheap test to perform and one for which animal welfare is not a significant issue. Animal welfare concerns can also be reduced for the acute oral toxicity test by careful consideration of the new substance's likely toxicity and use of the alternatives to the LD50 test that are now available.

By conducting the acute oral and Ames tests on the new substance the acceptability of the read-across can be further assessed when the results are known, i.e. if toxicity differences are found between the new substance and the structural analogue then further testing for other endpoints may be appropriate.

3.2.2. New tests should not be conducted within a read-across in an attempt to remove an unwanted classification.

Read-across applications occasionally request that certain data be read-across to the new substance from the structural analogue but ask to perform certain tests to remove an undesirable classification, such as that for skin sensitisation. In such cases, if certain tests were performed and gave a different result to the analogue, then the substance similarities and the whole scientific basis for read-across of other data would need to be questioned.

3.2.3. Within a series of structurally similar new substances, the two substances at either end of the series may be the only two that need to be fully tested.

This is a concept similar to the family approach already used in the notification of polymers and is based on the assumption that, in principle, members of a series of similar substances will possess similar types of hazard. The concept extends to the assumption that, due to higher solubility, low molecular weight members of a series of similar substances will be likely to produce greater toxicological effects than higher molecular weight members.

One read-across request received by HSE involved a structurally similar series of 4 substances differing in carbon numbers. The result was full base-set testing of the low molecular weight member of the series and limited testing on the highest molecular weight member. For the other group members all toxicological data for base-set notification was read-across (spectra and analysis data being provided).

3.2.4. For regulatory purposes it is often more straightforward to read-across positive findings.

Negative results obtained on structural analogues, by outdated methods for example, are always questioned. A positive result obtained in a similar way would still be carefully assessed but would be more likely to be accepted, as its read-across would be adopting a precautionary position with respect to the protection of human health.

APPENDIX 3

A strategy for assessing the suitability of environmental read-across data under a chemical notification scheme (Hanway, 2002b)

Dr Ros Hanway, Industrial Chemicals Unit, Health and Safety Executive, Magdalen House, Stanley Precinct, Bootle, Merseyside, L20 3QZ, UK.

The needs and principles that require consideration are outlined below.

2.1. Comparison of the chemical structure, including the purity and impurity profiles

The similarity of the chemical structure, and the purity and impurity profiles of the new substance and the structural analogue needs to be assessed. The fundamental basis for any read-across decision must be that the chemical structures of the analogues are sufficiently close for there to be a reasonable expectation of similar effects. The more divergent the structures, the lower will be confidence in making such a prediction. In general, where biologically active functional groups are present, they should be present in both structures and be in the same structural orientation so that any biological activity would be unaffected. There is a need to ensure that there are no differences in the purity or impurities that are likely to influence the overall toxicity.

2.2. Evaluation of physicochemical properties

In general, the UK c.a. would not accept read-across of the basic physicochemical properties. Such data provide the key descriptors of the substance, including verification that the substance is what it is said to be. Knowledge of these properties is also considered essential in decision making on read-across of other data between chemicals. Comparison of the physicochemical properties of the new substance and its related analogue, particularly the physical form, molecular weight, water solubility, partition coefficient, dissociation constants, hydrolytic stability and vapour pressure, provides useful information on their similarity, and also on their anticipated behaviour in a test system and test organism.

Changes in water solubility can clearly have a significant influence on the measurable ecotoxicity in aq uatic toxicity testing. In general, the greater the partitioning to octanol (higher log Kow), the greater will be the uptake from water to aquatic organisms. However, this is complicated by a general lowering of the water solubility for these same substances. Knowledge of these properties is thus clearly essential before decisions can be made on the comparability of toxicity.

2.3. Reading-across results obtained from outdated test methods.

Experience of using Annex V test methods/OECD test guidelines can lead to certain revisions to their protocols. If read-across data have not been produced using the most current Annex V test methods, particularly careful consideration of the quality and suitability of a method is important. Nevertheless, a test conducted to a non-OECD/Annex V guideline or even a non-GLP study, should not automatically be rejected. The reviewer must be satisfied, however, that a repeat study would yield the same results, and thus considerable expert judgement must be applied.

2.4. Use of QSARs

For aquatic toxicity, a number of QSARs exist that may help determine the likely toxicity of structural analogues, and these should be used where possible to inform decisions on both the extent and type of additional testing that might be required. In a series of structurally similar compounds, for example, the QSAR may indicate that the estimated toxicity of a substance may exceed the measured water solubility, indicating that no toxicity would actually be observed. Thus read-across from data on that substance, or to that substance from others in the series, would be open to question. The QSAR used should always be appropriate to the chemical class of compounds in question, and valid across the domain of the physicochemical properties of the chemical. The ability of QSARs to predict the outcome of tests which have been performed on related substances can be extremely informative in assessing whether or not QSAR predictions for the substance to be notified (and for which read-across of data is req uired) are of sufficient reliability to be considered as appropriate for making informed decisions.

QSARs may be able to provide a reasonable estimate of the “base line” toxicity of a substance, due to apolar narcosis. In some cases chemicals may have a more specific mode of action, resulting in greater toxicity than would be expected from apolar narcosis. This possibility needs to be considered and may affect the level of confidence in QSAR predictions.if a specific mode of toxicity is suspected fr a substance expert judgement will be required in assessing the reliability of QSAR predictions.

2.5. Basic toxicity tests to underpin the validity of read-across

Even where read-across can be agreed between two close structural analogues, some uncertainty will always remain. The UK c.a. has usually sought to address this uncertainty through the provision of some basic testing to underpin assumptions implicit in agreeing a read-across. Experience suggests that tests such as the acute toxicity to Daphnia on a new chemical can provide additional confidence that read-across of other data is possible, i.e. if toxicity differences are found between the new substance and the structural analogue then further testing for other endpoints may be appropriate. The acute Daphnia toxicity test raises few animal welfare issues while providing good confirmation of the comparability of aquatic toxicity.

2.6. Additional testing to remove an unwanted classification

Read-across applications occasionally request that certain data be read-across to the new substance from the structural analogue but ask to perform certain tests to remove an undesirable classification. In such cases, if certain tests were performed and gave a different result to the analogue, then the substance similarities and the whole scientific basis for read-across of other data would need to be questioned. The reviewer must therefore decide that either the new test is unnecessary since the available evidence would predict a similar result, or that read-across is not possible.

2.7. Testing representative members of a series of substances

This is a concept similar to the family approach already used in the notification of polymers and is based on the assumption that, in principle, members of a series of similar substances will possess similar types of hazard. Generally substances at either end of a series would be fully tested. The

series may span a range of water solubility, molecular weight, carbon chain length or other properties (several properties may co-vary). For example, from a series of substances the most water soluble and the least water solube substances may be tested as representatives of the series. For different groups of substances different physicochemical properties may be considered as appropriate indicators of the extremes of the group.

For aquatic toxicity, the measurement of effects is heavily dependent, not just on the water solubility, but also on the rate and extent to which a chemical can be taken up by the organism from the water. It cannot be assumed that the lowest molecular weight and/or size exert the highest toxicity, (although substances with a molecular mass >700- 800 often show low toxicity since they may not pass biological membranes due to steric hindrance) since the log Kow – a key indicator of uptake - is likely to rise with increased size. The acute Daphnia test on key members of a series may indicate whether one or more in the series should be more fully tested and act as a representative.

2.8. Read-across of conservative findings

For environmental degradation it is more precautionary to consider a group of substances as not readily biodegradable. In the case of a group of compounds where one end member is readily biodegradable and the other end member is not readily biodegradable additional testing would be required in order to determine whether of not any of the intermediate substances could be classified as readily biodegradable.

2.9. Higher tonnage testing programmes

New chemicals notification procedures are built around the principle of proportionality, ie that information requirements, including testing requirements, are proportionate to the level of supply and perceived risk. A read-across agreed at the base- set level must, therefore, be reassessed at higher tonnage triggers to determine whether the conclusions are still valid for any additional testing that might be sought. It must also be considered whether the extra supply tonnage would lead to a need to confirm all, or some of the base-set properties.