（参考資料６）

Guidance on Grouping of Chemicals, Third Edition

Series on Testing and Assessment

OECD

OECD Series on Testing Assessment
No. 418

Guidance on Grouping of Chemicals, Third Edition

This work was approved and declassified by the Chemicals and Biotechnology Committee on 23/10/2025.

Please cite this publication as:

OECD (2025), *Guidance on Grouping of Chemicals, Third Edition*, OECD Series on Testing and Assessment, No. 418, OECD Publishing, Paris, https://doi.org/10.1787/b254a158-en.

Photo credits: Egorov Artem/Shutterstock.com

About the OECD

The Organisation for Economic Co-operation and Development (OECD) is an intergovernmental organisation in which representatives of 38 countries in North and South America, Europe and the Asia and Pacific region, as well as the European Union, meet to co-ordinate and harmonise policies, discuss issues of mutual concern, and work together to respond to international problems. Most of the OECD's work is carried out by more than 200 specialised committees and working groups composed of member country delegates. Observers from several Partner countries and from interested international organisations attend many of the OECD's workshops and other meetings. Committees and working groups are served by the OECD Secretariat, located in Paris, France, which is organised into directorates and divisions.

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

Foreword

This document is part of OECD efforts to provide guidance for assessing the hazards of groups of chemicals, thus gaining efficiencies in the number of chemicals assessed and reducing reliance on animal testing. This guidance document reflects on the elements common to a number of different grouping approaches and real examples of the application of guidance to provide users with an understanding of basic concepts.

The publication of this guidance document is intended to improve the understanding and use of grouping approaches for assessing chemical hazards. Since techniques for assessing groups of substances are an evolving science, this third edition has been revised to reflect the state of the science. The guidance describes grouping approaches and workflows and provides illustrative cases that have been developed by stakeholders. Generally, the approaches described consider closely related chemicals as a group or category, rather than as individual chemicals. Using a category approach, not every chemical needs to be tested for every endpoint. Rather, the compilation of data for chemicals included in the category should be adequate to support a hazard assessment for the endpoints of interest and the compiled data should facilitate an estimate of hazard for the untested chemicals in the category.

This is the third edition of the OECD Guidance on the Grouping of Chemicals, with the first edition initially published in 2007. The third edition introduces new or revised guidance on the utility of New Approach Methodologies/Methods (NAMs) in developing groups and substantiating similarity. For example, updated guidance is provided on the use of omics technologies (e.g. metabolomics and transcriptomics data), phenotypic profiling, high throughput and high content screening (HTS/HCS) data, and (quantitative) structure-activity relationships ((Q)SARs). This guidance also discusses grouping approaches in the context of Integrated Approaches to Testing and Assessment (IATA) and Defined Approaches (DA) for specific endpoints, as well as new approaches to quantify read-across performance and uncertainties. The third edition also expands guidance on the use of mechanistic approaches, adverse outcome pathways (AOPs), and grouping of nanomaterials.

The third edition of the OECD Guidance on the Grouping of Chemicals was developed by a Steering Group of subject matter experts under the OECD Working Party on Hazard Assessment (WPHA). The update began in 2021, with experts from US Environmental Protection Agency (EPA) and the European Chemicals Agency (ECHA) generously agreeing to co-chair this effort beginning in 2023, with organisational support from a consultant provided in Q3/4 2023. The complete draft guidance was circulated for two rounds of review and written comment to the WPHA in March 2024 and January 2025. The complete draft guidance was circulated to the WPHA for two rounds of review and written comment, in March 2024 and January 2025. The WPHA approved the revised final draft by written procedure in August 2025.

This document is published under the responsibility of the Chemicals and Biotechnology Committee. The expectation is that this guidance will evolve continuously, based on experiences among OECD stakeholders and across regulatory frameworks.

Table of contents

About the OECD 3

Foreword 4

List of selected abbreviations 8

1 Introduction 14
1.1. Introduction 14

2 Explanation of the Grouping Approaches 17
2.1. Introduction and concepts 17
2.2. Considerations when grouping chemicals 25
2.3. Selecting analogues/Creating chemical categories and setting boundaries 28
2.4. Evaluating analogue and chemical category approaches with other similarity context information 32
2.5. Uncertainty analysis 39
2.6. Chapter 2 Annex - Examples 45

3 Techniques or Methods for Data Gap Filling 53
3.1. Introduction 53
3.2. Read-across 54
3.3. Trend analysis and computational methods based on internal models 86
3.4. Computational methods based on external models 89
3.5. Reporting on incremental changes, trend analysis and computational methods 90

4 Analogue Approach: A Stepwise Procedure for Identifying Analogues and Read-across 92
4.1. Introduction 92
4.2. Stepwise approach to read-across using the analogue approach 94

5 Category Approach: A Stepwise Procedure for Grouping Chemicals and Read-across 108
5.1. Introduction 108
5.2. Stepwise approach to the formation of chemical categories 111
5.3. Category development using an AOP 121

6 Guidance on Specific Types of Categories 123
6.1. Chain length 123

6.2. Isomers 124
6.3. Ionisable compounds 126
6.4. Metabolic or degradation pathways and toxicokinetics 127
6.5. Substances with more than one constituent and mixtures 132
6.6. Unknown or Variable composition, Complex reaction products or Biological material (UVCB substances, or UVCBs) 136
6.7. Metals and inorganic compounds 148
6.8. Grouping of nanomaterials 173

7 Reporting Formats for Analogue and Category Evaluations 186
Reporting format for analogue approach 187
Reporting format for chemical categories 197

References 209

Appendix 242
Computational tools 242
Grouping and read-across reporting template 259
Data matrix template 260
Reporting of omics data for chemical grouping 260

Glossary of selected terms 261

FIGURES

Figure 1. A multi-step workflow for chemical grouping based on bioactivity similarity and a plausible toxicological interpretation of the molecular formula 22
Figure 2. Graphical representation of a chemical category and some approaches for filling data gaps 25
Figure 3. Main aspects to consider for uncertainty assessment 41
Figure 4. Stepwise approach to an analogue approach 95
Figure 5. Stepwise approach to category development 110

TABLES

Table 1. Options for introducing bioactivity similarity as a line of evidence in a grouping hypothesis 35
Table 2. Key components of uncertainty related to the similarity rationale 43
Table 3. Aspects related to uncertainties covered by different frameworks 45
Table 4. Illustrative examples of quantitative and qualitative read-across 60
Table 5. Examples of categories based on constituent/component analysis 65
Table 6. Example for characterising category composition of UVCBs 66
Table 7. Examples of structures and functional groups in selected categories from the OECD HPV Programme and IATA Case Study Project 68
Table 8. Physicochemical properties important for hazard assessment 74
Table 9. Examples of categories with similarities in ADME/bioavailability 79
Table 10. Examples of categories with similarities in MOA 86
Table 11. Examples of structural differences of analogues discussed in the OECD IATA Case Study Project (OECD, 2016b) 100
Table 12. Examples of factors related to structural differences that may impact on toxicity (OECD, 2017a) 101
Table 13. Selected example tools for analogue searching 107
Table 14. Qualitative use of mechanistic information in forming chemical categories 122
Table 15. Toxic Equivalency Factors (TEFs) for `four dioxins 135
Table 16. Matrix compiling all lines of evidence with listing of examples of relevant types of data (other relevant types can be added in additional rows) 152

Table 17. Example of matrix compiling all lines of evidence for human health endpoints with illustration of relevant types of data (other relevant types to be added in additional rows) 153
Table 18. Matrix compiling all lines of evidence for the environmental compartment with listing of some relevant types of data (other relevant types to be added in additional rows) 156
Table 19. Framework for grouping substances in the environment, based on (ECHA, 2017b) 159
Table 20. Matrix compiling all data sources available for the grouping of Co substances 163
Table 21. Chronic aquatic nickel toxicity for three taxa as a function of the counter-ion in the exposure. The chronic toxicity shown within this table includes reported NOECs, LOECs and EC_{10}s 170
Table 22. Key values for chemical safety assessments as reported in the substance EU REACH registration dossier for freshwater invertebrates 171
Table 23. Ionic concentrations in European natural waters 171
Table 24. Summary of quantitative and qualitative lines of evidence supporting the grouping and read-across for nickel ions in the environment 172
Table 25. Examples of case studies that have employed the GRACIOUS Framework by (Stone et al., 2020) to group nanomaterials for either human or environmental hazards 185
Table 26. Example data matrix, analogue approach 190
Table 27. Example of uncertainty reporting related to the similarity rationale 195
Table 28. Examples for structural matrix 199
Table 29. Example data matrix, chemical category 202
Table 30. Example of uncertainty reporting related to the similarity rationale 206
Table 31. Computational tools supporting the development of the categories – related to specific aspects of different endpoints 242
Table 32. References and context of use for prediction tools 247

BOXES

Box 2.1. Chapter 2 summary 17
Box 3.1. Chapter 3 summary 53
Box 4.1. Chapter 4 summary 92
Box 4.2. Equation 1. Tanimoto index 97
Box 5.1. Chapter 5 summary 108
Box 6.1. Chapter 6 summary 123
Box 6.2. Equation 2. Calculating TEF 134
Box 6.3. Equation 3. Calculating TEQ 135
Box 6.4. Equation 4. Calculation of Toxic Units (TU) 158
Box 7.1. Chapter 7 summary 186

List of selected abbreviations

ADME	Absorption, Distribution, Metabolism, Excretion
AE	Assessment Element (in relation to the RAAF)
AIM	Analog Identification Methodology
AOP	Adverse Outcome Pathway
AR	Androgen Receptor
ARN	Assessment of Regulatory Needs
BAF	Bioaccumulation Factor
BCF	Bioconcentration Factor
BfR	German Federal Institute for Risk Assessment (*Bundesinstitut für Risikobewertung*)
CAS	Chemical Abstracts Service
CG-ARM	OECD Chemical Grouping Application Reporting Module
CLP	Classification, Labelling and Packaging Regulation (EU)
CMP	Chemicals Management Plan (Canada)
CONCAWE	Conservation of Clean Air and Water in Europe: The oil companies' European Organisation for Environment, Health and Safety in Refining and Distribution
CRED	Criteria for Reporting and evaluating Ecotoxicity Data
DA	Defined Approach
DTU	Technical University of Denmark
EC	European Commission
ECETOC	European Centre for Ecotoxicology and Toxicology Of Chemicals
ECHA	European Chemicals Agency
EFSA	European Food Safety Authority
EPA	Environmental Protection Agency (US or Danish)

ER	Estrogen Receptor
EU	European Union
EURL ECVAM	EU Reference Laboratory for alternatives to animal testing
GenRA	Generalised Read-Across
GHS	Globally Harmonised System (for the classification of chemicals)
HCS	High Content Screening
HPV	High Production Volume
HSPA	Hydrocarbon Solvents Producers Association
HTPP	High Throughput Phenotypic Profiling
HTS	High Throughput Screening
HTTK	High Throughput Toxicokinetics
IATA	Integrated Approaches to Testing and Assessment
ICE	Integrated Chemical Environment
ICPS	International Classification for Patient Safety
IHSC	International Hydrocarbon Solvents Consortium
ITS	Intelligent (Integrated) Testing Strategy
IUCLID	International Uniform Chemical Information Database
IUPAC	International Union of Pure and Applied Chemistry
KB	Knowledge Base
KE	Key Event
K_{oc}	Organic-carbon partition coefficient
K_{ow}	Octanol-water partition coefficient
log K_{oc}	log of the organic-carbon partition coefficient
LC_{50}	(Lethal Concentration, 50%) concentration of a compound that causes 50% lethality of the animals in a test batch
LED_{05}	Lower 95% confidence limit of the effective dose
LD_{50}	(Lethal Dose, 50%) dose of a compound that causes 50% lethality of the animals in a test batch

LOAEL	Lowest Observed Adverse Effect Level
LOEC	Lowest Observed Effect Concentration
LOEL	Lowest Observed Effect Level
MIE	Molecular Initiating Event
MOA	Mode and/or Mechanism Of Action
MW	Molecular Weight
NAFTA	North American Free Trade Agreement
NAM	New Approach Methodology/Method
NCS	Natural Complex Substances
NEF	Neurotoxic Equivalent Factor
NICA	Constituent distribution and relative presence in hydrocarbon solvents
NIOSH	National Institute for Occupational Safety & Health (US)
NM	NanoMaterials
NOAEC	No Observed Adverse Effect Concentration
NOAEL	No Observed Adverse Effect Level
NOEC	No Observed Effect Concentration
NOEL	No Observed Effect Level
NTA	Non-Targeted Analysis
OECD	Organisation for Economic Co-operation and Development
OORF	OECD Omics Reporting Framework
OPERA	OPEn structure-activity/property Relationship App
PAH	PolyAromatic Hydrocarbon
PBT	Persistent, Bioaccumulative and Toxic
PBD	Physiologically Based Dynamic
PBK	Physiologically Based Kinetic
PCBs	PolyChlorinated Biphenyls

PFAS	Per- and polyFluoroAlkyl Substances
PNEC	Predicted No Effect Concentration
PPRTV	Provisional Peer Reviewed Toxicity Values
POD	Point of Departure
QAAR	Quantitative Activity-Activity Relationship
QAF	(Q)SAR Assessment Framework
QMRF	(Q)SAR Model Reporting Format
QPRF	(Q)SAR Prediction Reporting Format
(Q)SAR	(Quantitative) Structure-Activity Relationships
RAAF	Read-across Assessment Framework (as developed by ECHA)
REACH	Registration, Evaluation, Authorisation and Restriction of Chemicals (EU)
RPF	Relative Potency Factor
SAR	Structure Activity Relationship
SciRAP	Science in Risk Assessment and Policy
SEURAT	Safety Evaluation Ultimately Replacing Animal Testing (EU Framework Project)
SIDS	Screening Information Data Set (OECD)
SMILES	Simplified Molecular Input Line Entry System
SMARTS	SMILES arbitrary target specification
TEF	Toxic Equivalency Factor
TEQ	Toxic Equivalents (Approach)
TEST	Toxicity Estimation Software Tool
TG	Test Guideline (in relation to OECD)
TK	Toxicokinetics
TOXNET	Toxicology Data Network
TSCA	Toxic Substances Control Act (US)
ToxCast	Toxicity Forecaster programme (US)

UN	United Nations
UNEP	United Nations Environment Programme
US	United States (of America)
UVCB	Substances of Unknown or Variable composition, Complex reaction products and Biological materials
vPvB	Very Persistent and Very Bioaccumulative
WHO	World Health Organization
WoE	Weight of Evidence
WPHA	OECD Working Party on Hazard Assessment
WPMN	OECD Working Party on Manufactured Nanomaterials

Executive Summary

This document provides guidance for regulators reviewing the suitability of using grouping data to inform regulatory decisions, as well as for scientists (and/or regulatory registrants) generating research data in the grouping of chemicals. Due to the evolution of grouping methodologies and tools, and the inherent complexity of this topic, the intent of this guidance is not to provide an opinion on the regulatory acceptability of particular grouping approaches. Early consultations between industry and authorities are recommended, where possible, to ensure that a particular grouping approach is compatible with regulatory requirements. Users should note that the OECD guidance cannot cover all possible decision contexts or regulatory requirements, nor is this guidance prescriptive regarding which grouping approaches are best suited for regulatory decisions or accepted by different regulatory authorities.

Approaches for chemical grouping have been undertaken on an ad hoc basis in regulatory programmes for many years. Indeed, guidance was first developed by the US Environmental Protection Agency (EPA) in support of the US High Production Volume (HPV) Challenge Program in 1998. The same guidance was also embedded into the “OECD Manual for the Assessment of Chemicals”. Since then, guidance has evolved based on grouping approaches included in regulatory frameworks and voluntary hazard assessment frameworks, as well as experiences shared with the OECD Chemicals Programme.

Chapter 1 provides an overview and introduction to chemical grouping as a method to reduce resources required to assess chemical hazards. The chapter introduces key concepts and includes a brief history of grouping approaches used in a regulatory decision-making context. Chapter 2 outlines general aspects of grouping chemicals such as the identification of analogues or members of categories, the biological and/or mechanistic basis, and the robustness of analogue or chemical category approaches. Chapter 2 also provides guidance on the use of (Q)SARs for data evaluation, read-across hypothesis generation and data gap filling, and explains where these components fit in the context of an IATA or DA. Chapter 2 concludes with a discussion of the challenges of assessing and addressing residual uncertainties within analogue and category approaches, and provides a comparison of approaches, and examples. Chapter 3 explains read-across (including for continuous endpoints and trend analysis) and (Q)SARs approaches used for data gap filling and provides details of computational tools that can facilitate grouping procedures outlined in Chapters 4 and 5.

While Chapters 2 and 3 provide explanations on the scientific and methodological background of the analogue and category approaches, Chapters 4 to 7 focus on practical aspects for forming and documenting analogue and chemical category approaches.

Chapters 4 and 5 provide guidance on stepwise procedures for analogue and chemical category approaches, respectively. Users may want to use this guidance document in a “modular” fashion, and therefore, chapters 4 and 5 include repeated text so they can be consulted independently.

Chapter 6 elaborates on considerations for grouping of chemicals based on chemistry, such as isomerism or metabolism, as well as for metals and inorganic compounds, substances of unknown or variable composition, complex reaction products or biological material (UVCBs), and nanomaterials. Finally, Chapter 7 proposes formats for documenting analogue and category approaches, which include uncertainty assessment frameworks discussed in Chapter 2.

1 Introduction

1.1. Introduction

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

For reasons of resources and animal welfare, it is important to reduce *in vivo* testing, where scientifically justifiable. The practice of predicting properties of chemicals is well established in regulatory science, and techniques to predict chemical properties are evolving with the development and application of scientific knowledge. One approach is to consider closely related chemicals as a group rather than as individual chemicals. If grouping is applied, not every chemical needs to be tested for every required endpoint. Rather, the data for chemicals and endpoints that have been tested can be used to estimate the corresponding properties for the untested chemicals and endpoints. Approaches for grouping chemicals to predict properties of some members of the group based on information available for other members of the group (i.e. read-across) depend on the purpose of the prediction (e.g. for commercial decision-making, screening and priority-setting of chemicals for further evaluation, filling information requirements in different regulatory schemes, hazard identification for classification and labelling, or use in risk assessment). Therefore, the level of resources, need for additional data generation, the extent of scientific justification, and scientific confidence in the associated assessment vary depending on the purpose of the prediction, the problem formulation, and the regulatory requirements. For example, the level of acceptable residual uncertainty tolerated for a screening and priority-setting scenario for thousands of chemicals may be greater than that allowed for their risk assessment. This guidance document cannot cover all purposes and regulatory requirements that may apply, nor is it prescriptive, as far as which grouping approaches would or would not be acceptable by different regulatory bodies. The OECD Guidance on Grouping of Chemicals aims to encompass different approaches and interpretations and reflect on the elements common to a number of applications.

Grouping of chemicals can lead to the application of a category or an analogue approach. In the analogue approach, where comparisons are made between a very limited number of chemicals, endpoint information for one chemical (the “source”) is used to predict the same endpoint for another chemical, (the “target”), which is considered to be “similar” usually on the basis of structural similarity and similar properties and/or activities. In the category approach, where comparisons are made between more chemicals, chemicals whose physicochemical, toxicological and ecotoxicological properties are similar or follow a regular pattern as a result of structural similarity may be considered as a group, or “category” of chemicals. In this approach, the properties of the individual members of a category are assessed on the basis of information for a given endpoint available for all category members.

Chemical grouping and read-across is commonly used as an approach to fill data gaps for individual chemicals and avoid the need to fill gaps by extensive testing. A data gap is a physicochemical, environmental fate, ecotoxicological, or mammalian toxicological/human health endpoint for which data

are not available when required for an assessment. “Data gap filling” is the process of providing data to inform upon a particular endpoint by whatever means is scientifically justified including alternative techniques to direct testing. Read-across and trend analysis are methods that may be used for data gap filling as described in this document.

The first edition of this document was developed based on existing cases involving chemical categories assessed within the OECD Cooperative Chemicals Assessment Programme[1] (formerly the OECD High Production Volume (HPV) Chemicals Programme), the US HPV Challenge Program[2], other US EPA programmes, the EU Existing Substances Programme (replaced by EU REACH[3] (Registration, Evaluation, Authorisation and Restriction of Chemicals) in 2006), the EU activity on classification and labelling[4], Canada’s Chemicals Management Plan[5] (CMP), and the experience gained from the OECD Workshop on the development and use of chemical categories held in 2004. A major milestone from the first edition was to clarify the interplay between (Q)SARs and categories and to provide some standardisation in terminology when referring to the grouping approach being used versus data gap filling.

The 2014 second edition included considerations on understanding of mechanistic interactions between the chemical of interest and the biological target, and key events leading to adverse effects (i.e. the Adverse Outcome Pathway [AOP] concept). This second edition included updates reflecting experiences gained from the OECD Workshop on Using Mechanistic Information in Forming Chemical Categories (OECD, 2011a), as well as insights following EU REACH legislation entering into force (ECETOC, 2012; Patlewicz et al., 2013a). The second edition also addressed the formation of categories for test plan and hazard assessment purposes. It also provided guidance reporting data (e.g. data matrices of all available data for category members, with an indication of the data gap filling technique proposed) and some challenges identifying and addressing uncertainties.

This 2025 third edition includes updates based on the experiences of the IATA Case Studies Project and Working Party on Manufactured Nanomaterials (WPMN). The updates also include lessons from the OECD IATA Case Studies and other related efforts (e.g. Safety Evaluation Ultimately Replacing Animal Testing (SEURAT)-1, EU ToxRisk, Accelerating the Pace of Chemical Risk Assessment (APCRA) Project, US EPA Provisional Peer Reviewed Toxicity Values (PPRTV) assessments, the US EPA National Testing Strategy for per- and polyfluoroalkyl substances (PFAS)). Experience with the use of read-across in the context of EU REACH has been further elaborated to include reference and experience applying ECHA’s Read-Across Assessment Framework (RAAF; first published in 2015 and extended in 2017) which provides a framework and guidance for consistent evaluation of read-across approaches (ECHA, 2017b). This third edition of the OECD Guidance on Grouping of Chemicals also introduces new or revised guidance on the

[1] OECD Cooperative Chemicals Assessment Programme (CoCAP)

https://web-archive.oecd.org/2016-10-19/58206-cocap-cooperative-chemicals-assessment-programme.htm

[2] US EPA High Production Volume (HPV) Challenge Program
https://nepis.epa.gov/Exe/ZyPURL.cgi?Dockey=P1011R0L.txt.

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

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

[5] Canada – Chemicals Management Plan (CMP) https://www.canada.ca/en/health-canada/services/chemical-substances/chemicals-management-plan.html

utility of NAMs in providing mechanistic information to support developing groups or substantiating similarity (e.g. bioactivity similarity). In the context of this guidance, these discussed NAMs include *in vitro* assays, omics technologies (e.g. transcriptomics, metabolomics, proteomics data), phenotypic profiling data, HTS/HCS data, (Q)SARs and other computational models. This guidance provides additional information on grouping approaches in the context of IATA (including DAs) for specific endpoints (OECD, 2020a), on describing uncertainties, and quantification of read-across performance. This third edition also introduces new guidance on grouping of nanomaterials.

Guidance on the regulatory application of (Q)SAR methods for providing data for specific endpoints is outside of the scope of this document and can for example be found in the following documents:

- Section 3.3 of the OECD Manual for the Assessment of Chemicals provides guidance on the use of SAR in the HPV Chemicals Programme (OECD, 2000a).
- OECD Report on the Regulatory Uses and Applications in OECD Member Countries of (Q)SAR Models in the Assessment of New and Existing Chemicals (OECD, 2006a) summarises the experience of OECD Member Countries with (Q)SAR applications.
- OECD report on the principles for the validation, for regulatory purposes, of (Q)SAR models (OECD, 2004a) and an accompanying OECD guidance document (OECD, 2014c).
- Report of the Workshop on Structural Alerts for the OECD (Q)SAR Application Toolbox (OECD, 2009b).
- Training for the QSAR Toolbox[6].
- NAFTA (2012). Technical working group on pesticides (TWG) (Quantitative) structure activity relationship [(Q)SAR] guidance Document[7].
- ECETOC (2012) ECETOC Technical Report 116: Category approaches, read-across, (Q)SAR.
- Recent Advances in QSAR Studies: Methods and Applications (2010) Edited by T Puzyn, J Leszczynski, MTD Cronin (Enoch, 2010).
- (Q)SAR Assessment Framework: Guidance for the regulatory assessment of (Quantitative) Structure Activity Relationship models, predictions, and results based on multiple predictions, 2nd edition (OECD, 2024a).

[6] https://qsartoolbox.org/support/

[7] https://archive.epa.gov/pesticides/news/web/html/qsar-guidance.html

2 Explanation of the Grouping Approaches

Box 2.1. Chapter 2 summary

This chapter:

- Explains what a category is and outlines relevant concepts.
- Outlines general aspects of grouping chemicals such as the identification of analogues or members of categories, the biological and/or mechanistic basis for using analogues or chemical categories, and the robustness of both approaches.
- Describes the close relationship that exists between (Quantitative) Structure-Activity Relationships ((Q)SARs) and categories, both in terms of the concepts and the use of (Q)SARs for data evaluation, read-across hypothesis generation and data gap filling, and where these components fit in the context of an Integrated Approach for Testing and Assessment (IATA) or a Defined Approach (DA).
- Discusses the challenges of assessing and addressing residual uncertainties within analogue and category approaches.

2.1. Introduction and concepts

In this OECD guidance document, the term 'grouping' or 'chemical grouping' describes the general approach for considering more than one chemical at the same time. It can lead to formation of a chemical category or identification of chemical analogues with the aim of filling data gaps[8] as appropriate. The category or the analogue approach makes it possible to extend the use of empirical data to similar untested chemicals, so that reliable estimates that are adequate for classification and labelling and/or risk assessment can be made without further testing. In this way, both approaches are important since they provide an alternative to testing individual chemicals which should lead to a decrease in the use of animal testing. In addition, they will increase the knowledge of the hazard properties of chemicals that may otherwise remain untested and provide for an increased level of protection for human health and the environment.

[8] A data gap is a physicochemical, environmental fate, ecotoxicological, or mammalian toxicological/human health endpoint for which data are not available when required for an assessment.

2.1.1. Analogue approach

When the focus of the assessment is on filling data gaps for one specific "target" chemical, empirical data from one or more similar chemicals, ("the analogues"[9] or "source" substances[10]) can be used to predict the same endpoint[11] for the "target" chemical[12], which is considered to be "similar" (see Glossary, "similarity").

Such an analogue approach is particularly compelling when the target and source chemicals share a known common mode (and/or mechanism) of action (MOA)[13,14], and associated adverse effects[15] resulting from this mode (and/or mechanism) of action are being evaluated. The analogue approach could also be used in the absence of effects, as well as when no specific MOA is expected and toxicokinetic behaviour is not expected to differ significantly. In such cases, more evidence[16], or more lines of evidence, may be needed to substantiate the assessment.

2.1.2. Category approach

Chemicals whose physicochemical, toxicological and/or ecotoxicological properties are likely to be similar or follow a regular pattern usually as a result of structural similarity may be considered as a group, or 'category', of chemicals.

The assessment of chemicals by a category approach differs from the approach of assessing them individually. The properties of the individual chemicals in a category are evaluated using all the available information for a given endpoint for all the category members. This process is broader than those based on empirical data for any one particular chemical on its own.

For category members that lack data for one or more endpoints, the data gap can be filled in a number of ways, including by read-across from other category members. Within a chemical category, the members are often related by a trend in an effect for a given endpoint, and a trend analysis[17] can be carried out through deriving a model based on the data from one or more endpoints for the members of the category.

---

[9] An analogue (or analogous substance) is one substance that has been identified as exhibiting similarity (see definition in the Glossary of selected terms) to another substance.

[10] A source substance (or source analogue) is a chemical that has been identified as an appropriate chemical for use in a read-across based on similarity to the target chemical and existence of relevant data.

[11] An endpoint represents any single or group of physicochemical, biological, or environmental property that can be measured/modelled. An endpoint could be determined by different experimental protocols and under different experimental conditions.

[12] A target chemical is a substance of interest for which data gaps exist that need to be addressed.

[13] A mode of action describes a biologically plausible sequence of key events at different levels of biological organisation, starting with the exposure to a chemical and leading to an observed (adverse) effect.

[14] A mechanism of action is a detailed molecular description of the mechanistic interaction through which a substance/molecule produces its effect.

[15] An adverse effect refers to the change in the morphology, physiology, growth, development, reproduction, or life span of an organism, system, or (sub)population that results in an impairment of functional capacity, an impairment of the capacity to compensate for additional stress, or an increase in susceptibility to other influences (ICPS, 2004).

[16] In the context of grouping chemicals, evidence refers to similarities in chemical data that is used to justify reading across from source to target chemical(s) and developing chemical categories.

[17] Trend analysis refers to a data-gap filling method for "quantitative endpoints" (e.g. 96h-LC_{50} for fish) if a number of analogues (at least 3) with experimental results are identified.

This is most feasible in cases where there is a single independent variable associated with the endpoint of interest.

An advantage of a category approach is that identification of patterns of effects within a category may increase confidence in the reliability of the data available for the individual chemicals in the category, compared to evaluation of data on a chemical-by-chemical basis. The overall robustness of a category prediction depends on the amount and distribution of available data (data density) across the category. For example, data may not sufficiently cover all structural variations and therefore would not support the identification of trends or allow for reliable predictions.

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

2.1.3. Rationale and justification of the grouping approach

The rationale underpinning the analogue or category approach may be based on one or more similarities including: structure, physicochemical properties, chemical reactivity profile, bioactivity, conventional[18] toxicological profile and Absorption, Distribution, Metabolism[19], Excretion/Toxicokinetics (ADME/TK).

Examples could include:

- Common functional group(s) (e.g. aldehyde, epoxide, ester, specific metal ion).
- An incremental and constant change across the category (e.g. a chain length category), often observed in physicochemical properties, e.g. boiling point range.
- Common substructures which are associated with specific reactive mechanisms (e.g. protein or DNA binding).
- Common constituents or chemical classes, similar carbon range numbers. This is frequently the case with complex substances[20] often known as “substances of unknown or variable composition, complex reaction products or biological material” (UVCB substances). The naming and identity of UVCBs may include details of the starting materials, production process and known chemical composition.

[18] Conventional toxicology refers to the traditional methods and approaches used to assess the potential toxicity of substances, particularly in contrast to new approach methods (NAMs).

[19] Metabolism is a linked series of chemical reactions in the body to convert a chemical (i.e. a xenobiotic) to either an inactive compound or to a more active compound for excretion from the body.

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

- The likelihood of common precursors and/or breakdown products via physical or biological processes that result in structurally similar chemicals. For example, the "metabolic pathway approach" of examining related chemicals such as acids, esters, and salts. One may also incorporate similarity in the biotransformation pathways themselves as well as the biotransformation products (or metabolites) formed.
- Bioactivity similarity, for example similarity between two or more omics[21] profiles (or signatures), where those profiles (or signatures) are measured in a defined test system following exposure to two or more test substances, often as part of a bridging study[22].
- Similarity in toxicological profile may refer to phenotypic responses from conventional toxicology studies.
- A common MOA or adverse outcome pathway (AOP).

The definition of a group typically starts with evaluating structural similarity followed by investigating other contexts of similarity, including bioactivity similarity, though this is not mandatory - specific requirements may apply depending on the decision and regulatory context. For every category, the structural elements that category members have in common need to be described, along with structural differences that may occur in the category. The structural differences may or may not affect the endpoint of interest. Differences that are not expected to affect the endpoint of interest, should be argued based on careful investigation in relation to the read-across hypothesis, and are termed "allowed differences". The set of required structural elements and allowed structural differences define the applicability domain (or boundaries) of the category and define which substances can be part of the category and which are not (category membership). The category could then be further limited by other considerations such as physicochemical properties and/or bioactivities. The extent to which this level of specificity should be described and documented depends on the stated purpose. For a category approach that is being developed to formulate initial pragmatic groupings from larger inventories for prioritisation purposes, e.g. US EPA's National per- and polyfluoroalkyl substances (PFAS) Testing Strategy (www.epa.gov/assessing-and-managing-chemicals-under-tsca/national-pfas-testing-strategy; Patlewicz et al., 2024a), such specificity is not necessarily expected nor practical to implement.

While a category may be in principle based on one or more of the rationales above, in practice, endpoint justifications and supporting information will be multifaceted. All pre-existing experimental or other (e.g. from the literature) evidence relevant to the category needs to be addressed. This could include but is not limited to, similar effects in low-tier studies (e.g. studies of short duration); availability of bridging studies that build support for the commonality but may not necessarily be endpoint specific (such as *in vitro* or *in vivo* studies, e.g. toxicokinetic data, metabolomics or transcriptomics data); evidence from computational models (e.g. (Q)SAR or molecular docking models that predict the endpoint being read across or other endpoints of relevance for the similarity assessment or hypothesis/justification, based on structural or biological features); common bioavailability[23], metabolism (empirical or simulated) and

[21] In the context of this guidance document, omics refers to technologies that are used to measure a broad range of molecular responses to chemical exposure. Widely used approaches include transcriptomics (study of expression of multiple genes) and metabolomics (study of levels of multiple endogenous metabolites and the biochemical processes in which they are involved in), within a cell, tissue, or organism.

[22] Bridging studies are defined as but not limited to studies conducted to show relevance or create a bridge between the substances included in an analogue or category approach to establish the similarity in properties, (eco)toxicological profile, and/or environmental fate and behaviour.

[23] Bioavailability is defined as the extent to which a substance is taken up by an organism and distributed to an area within the organism. It is dependent upon physicochemical properties of the substance, anatomy and physiology of the organism, pharmacokinetics, and route of exposure (United Nations, 2013). An alternative definition for bioavailability is the rate and extent to which a substance can be taken up by an organism and is available for

reactivity profiles (captured as structural alerts or empirically derived) can also support commonality. Data, including from computational models, demonstrating a common MOA or AOP can also be used (see Section 2.4.3). It is important that the quality of the supporting information is described, including how the information is relevant for the endpoint under consideration for the grouping and associated read-across or how the information strengthens the grouping justification.

2.1.4. Supporting information, other concepts and approaches related to grouping and read-across

Different types of information contribute to grouping and read-across approaches, including those which aim to substantiate the similarity and read-across rationale. Many of the approaches used to generate and/or analyse this information are often referred to as new approach methods (NAMs). NAMs include the best available science that have resulted from expanded research in biotechnology and data science, and are consistent with the reduction, refinement, and replacement of animal testing approaches. The acronym NAMs is not used to describe methods that are exclusively non-animal.

For the purposes of this document, a number of these approaches are described in brief in the context of their role in supporting grouping and read-across.

High-Throughput Screening (HTS), High-Content Screening (HCS), and omics technologies

Data from HTS assays (e.g. US EPA's ToxCast programme), HCS assays or omics technologies such as transcriptomics and metabolomics can also play a role in either identifying candidate source analogues or substantiating bioactivity similarity (a type of biological similarity). These approaches generate multidimensional assay results, often in the form of differential abundance analyses (i.e. changes in expression of gene transcripts, protein levels or endogenous metabolite levels). In such cases, chemical grouping can use the multi-step workflow shown in Figure 1, starting with a clear rationale for applying HTS, HCS and/or omics technologies. Following data generation, the calculation of a pairwise similarity metric between assay profiles for two substances presents one practical and quantitative means of characterising bioactivity similarity for analogue identification and evaluation; hence the approach is referred to as bioactivity-based grouping. The results should be compiled and summarised in a manner that provides a clear link to the read-across proposed, with explicit relevance to the endpoint under consideration, and address specific aspects of the grouping and read-across hypothesis, where feasible. This can include a statistical assessment of the bioactivity similarity and potentially a description of a plausible toxicological interpretation of the grouping. A more detailed introduction to the use of bioactivity similarity for providing lines of evidence for the grouping hypothesis is described in (Viant et al., 2024a) and presented in Section 2.4.2 and Appendix (Chemical Grouping Application Reporting Module [CG-ARM]).

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

Figure 1. A multi-step workflow for chemical grouping based on bioactivity similarity and a plausible toxicological interpretation of the molecular formula

(Quantitative) Structure-Activity Relationships ((Q)SARs)

Structure-activity relationships (SARs) and quantitative structure-activity relationships (QSARs) are collectively referred to as (Q)SARs. A SAR is a qualitative relationship that relates a (sub)structure to the presence or absence of a property or activity of interest. SARs can be helpful in identifying candidate source analogues and may support the qualitative evaluation of the analogues identified as members of a category. For example, SARs have been implemented as endpoint-specific alerts in some of the profilers in the QSAR Toolbox, such as the protein binding alerts for skin sensitisation, the DNA binding alerts for AMES mutagenicity, etc. Their use can aid in the construction of chemical categories to assess specific endpoints of interest (e.g. skin sensitisation, *in vitro* mutagenicity, etc.).

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

The OECD outlined key principles (OECD, 2004f; OECD, 2014c; OECD 2024a) for evaluating (Q)SAR models used for regulatory purposes. (Q)SAR models used for regulatory purposes should have adequate and reliable documentation that demonstrates the following principles:

1. A defined endpoint.
2. An unambiguous algorithm.
3. A defined domain of applicability.
4. Appropriate measures of goodness-of-fit robustness and predictivity.
5. A mechanistic interpretation, if possible.

In addition, the (Q)SAR Assessment Framework (QAF) (OECD, 2024a) establishes four principles for the assessment of (Q)SAR predictions and results from multiple predictions for regulatory purposes:

1. The model input(s) should be correct.
2. The substance should be within the applicability domain of the model.
3. The prediction(s) should be reliable.
4. The outcome should be fit for the regulatory purpose.

The importance of mechanistic understanding is twofold. First, the structure-activity relationships. i.e. linking a biological activity to a molecular (sub)structure, provide useful insights increasing the reliability and causality of the (Q)SAR model. Secondly, with mechanistic understanding how the effects relate to

the structure, a series of structural requirements can be described which define the mechanistic boundaries for a reliable applicability domain of the (Q)SAR model.

Similar to QSARs, quantitative activity-activity relationships (QAARs) are mathematical relationships between biological endpoints, which can occur in the same or different species. QAARs assume the MOA obtained for one endpoint is applicable to a similar endpoint in the same species or the "same" endpoint in a different species, since the underlying processes are the same (e.g. partitioning, reactivity, enzyme inhibition). Examples of QAARs include prediction of daphnid toxicity from *Tetrahymena pyriformis* or predicting acute oral toxicity from cytotoxicity measurements. In recent years, the concept of QAARs has broadened in scope; rather than a relationship between two endpoints in the same or different species, the concept has evolved to include the integration of different orthogonal assays in a computational network model and the use of assays as descriptors along with structural features. Illustrative examples include a model that integrated 18 *in vitro*, HTS assays measuring estrogen receptor (ER) binding, dimerisation, chromatin binding, transcriptional activation, and ER-dependent cell proliferation to predict ER agonists/antagonists (Judson et al., 2015); the use of cell viability assays in the development of carcinogenicity models (Zhu et al., 2008); and the use of HTS data in the development of acute rodent lethality (Sedykh et al., 2010).

Adverse Outcome Pathways (AOP)

An AOP delineates the documented, plausible, and testable (in principle) process by which a chemical induces molecular perturbations and the associated biological responses at the sub-cellular, cellular, tissue, organ, whole animal and population level. The AOP framework is based on the concept that toxicity results from a chemical interacting with an initial target (e.g. membrane, receptor) defined as the Molecular Initiating Event (MIE)[24]. Subsequently, a series of Key Events (KE)[25] that can be individually documented and tested are triggered, resulting in an adverse outcome (e.g. reproductive failure, neurotoxicity). Obviously, several pathways can share the same or interlinked KEs and/or the same adverse outcome, and each constitutes an individual AOP which can be linked in an AOP network.

The current AOP Knowledge Base (AOP KB) and associated tools (e.g. AOP wiki, Effectopedia) can be useful resources to identify relevant AOPs and related MIEs and KEs. AOP KB also contains AOPs under development. A list of OECD-endorsed AOPs can be found at OECD Series on Adverse Outcome Pathways[26].

Furthermore, the "OECD Guidance Document for Developing and Assessing Adverse Outcome Pathways" (OECD, 2017g) gives an overview of the vocabulary and concepts, as well as an insight into the development of an AOP, including identification and use of relevant scientific data and resulting knowledge. The complementary "Users' Handbook supplement to the Guidance Document for developing and assessing Adverse Outcome Pathways" (OECD, 2016a) provides more detailed information on AOP development as well as some guidance for the evaluation of confidence in the underlying information.

An OECD project launched by the Working Party on Manufactured Nanomaterials (WPMN) also investigated the development of AOPs for categorisation of the mechanisms of toxic action and risk assessment of nanomaterials. The outcome of the project includes a document describing a methodology

[24] A molecular initiating event (MIE) is a specialised type of key event that represents the initial point of chemical/stressor interaction at the molecular level within the organism that results in a perturbation that starts the AOP.

[25] A key event is a change in biological or physiological state that is both measurable and essential to the progression of a defined biological perturbation leading to a specific adverse outcome.

[26] https://www.oecd.org/en/publications/oecd-series-on-adverse-outcome-pathways_2415170x.html

to identify, analyse and evaluate existing nanotoxicology literature with the objective to prioritise KEs relevant for manufactured nanomaterials (OECD, 2020c).

How AOPs can mechanistically inform the grouping of chemicals is elaborated in Section 2.4.3.

Integrated Approaches to Testing and Assessment (IATA)

Integrated approaches to testing and assessments (IATAs) are based on multiple information sources used for hazard identification, hazard characterisation and/or safety assessment of chemicals (OECD, 2016e). An IATA integrates and weights relevant evidence and guides the targeted generation of new data where required, to inform regulatory decision-making regarding potential hazard and/or risk. Within an IATA, data from various information sources are evaluated and can include *in silico* models such as (Q)SARs, as well as grouping approaches. In essence, a grouping and read-across approach can be a component of the IATA (OECD, 2020a). Defined Approaches (DAs) are a type of IATA where all included information sources are fixed and are combined and interpreted using a fixed data interpretation procedure (e.g. a rules-based, algorithm). IATAs, DAs, and IATA-related components and concepts are described, along with an overview of more detailed guidance, in the "Overview of Concepts and Available Guidance related to Integrated Approaches to Testing and Assessment (IATA)" (OECD, 2020a).

2.1.5. Interpolation and extrapolation

While other data-gap filling approaches will be discussed in more detail in Chapter 3 it is worth mentioning interpolation and extrapolation.

Interpolation uses data from category members on either side of a data-poor category member to predict its hazards. Within a category, where trends in toxicity or factors influencing toxicity have been identified and the category members can be practically arranged in line with the trend as illustrated in Figure 2, one can interpolate. In contrast, extrapolation is the process where data from category members at one side of the category is used to predict the hazards of those members at the other side. An analogue approach is an extrapolation, unless analogues are identified that bracket the target chemical.

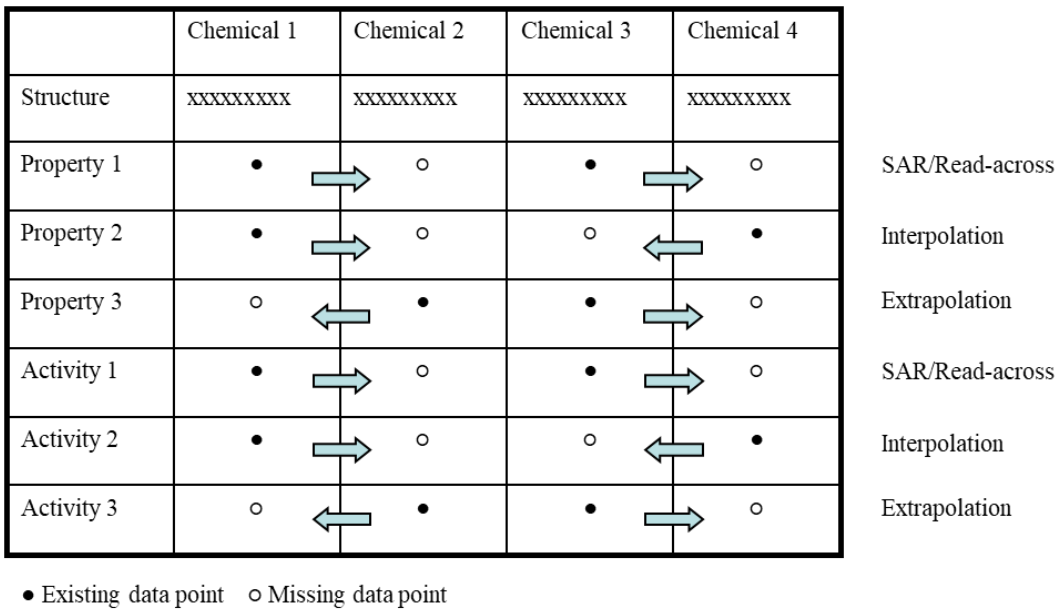
Extrapolation may represent a worst case or an underestimate of the toxicity and is generally perceived to be more uncertain (and therefore less reliable) than interpolation. The quality of the trend, among other factors, can influence the uncertainty, The establishment of a trend requires a dependent and independent variable, thus, the trend will depend on the choice and quality of the independent variable. If a trend is poorly defined or missing, interpolation and extrapolation approaches will be uncertain. therefore, it may seem logical for interpolation to be more 'acceptable' than extrapolation. However, the degree of uncertainty is not really due to the interpolation or extrapolation of data per se, but rather is more dependent on the robustness of the category rationale (ECETOC, 2012). Robustness is, in turn, dependent on the size of the category, the quantity and quality of data available for each category member, and the distribution of available data across the category (data density).

For large, data-rich categories, trends in toxicity are more likely to be characterised, such that data-gap filling using interpolation or extrapolation is more likely to be robust and useful. However, in cases where an analogue approach has been used, where a category consists of a small number of members, or where a large category comprises only a handful of members with relevant data, trends can be difficult to identify, and interpolation is often not possible. In these situations, extrapolation of data from one chemical to another may be the only possibility for filling the data gaps. To reduce uncertainty, a Weight of Evidence (WoE)[27] proposal can be developed that incorporates the use of supporting data. The OECD "Guiding Principles and Key Elements for Establishing a Weight of Evidence for Chemical Assessment" (OECD,

[27] WoE refers to a stepwise process/approach of collecting and weighing evidence to reach a conclusion on a particular problem formulation including assessment of the degree of confidence.

2019) was intended to help regulators to develop a consistent and transparent approach for delivery of evidence. Other schemes for WoE include the *in silico* toxicology protocols (Myatt et al., 2018; 2022) which describe how to document *in silico* and empirical data for specific endpoints. There will always be some degree of uncertainty (see Section 2.5.1), but this should include consideration of uncertainties associated with all test data, and not only uncertainties linked to category or analogue approaches (ECETOC, 2012, Pham et al., 2020).

Figure 2. Graphical representation of a chemical category and some approaches for filling data gaps



Note: Figure highlights the differences between the concepts of interpolation and extrapolation within a category approach. It also highlights the complementary use of external SAR models as an alternative qualitative data gap filling technique.

2.2. Considerations when grouping chemicals

The benefits of using of analogue and category approaches are outlined in detail below and include saving money, time, and animals when evaluating chemical safety. However, there are also scientific and practical challenges related to gaining access to quality data required for data-gap filling approaches, documenting the appropriate level of study information, and having the necessary information to characterise the target chemical and source analogues, as well as their respective impurity profiles. The generation of new data may be needed to substantiate the hypothesised basis of the category/analogue approach. In addition, evaluation of a read-across approach for regulatory acceptance may be more time-consuming than evaluating a single study for an endpoint. While these challenges maybe be associated with financial and human resource costs, these should be lower than costs associated with performing a full set of experimental studies to meet regulatory requirements.

If the only data gaps to be filled were for physicochemical or environmental fate endpoints, (Q)SARs may be more practical applied to predicting specific properties rather than using a full grouping approach. Obviously, this will depend on the endpoints for which there are data gaps and the robustness of the (Q)SAR models available for use. There are many freely available (Q)SAR models that have been developed, characterised with respect to the OECD Validation Principles, and documented in accordance

with the QAF (OECD, 2024a). Specific computational tools that have these encoded as models or their associated predictions are discussed in more detail in Chapter 3

The AOP (OECD, 2012a) concept can support building toxicologically meaningful categories by characterising the mechanism(s) by which chemicals lead to an adverse effect. Currently, this level of toxicological understanding needed to develop an AOP is difficult to achieve even using state-of-the-art HCS approaches (e.g. omics technologies). Biologically more complex endpoints are particularly challenging to address. However, in the absence of an AOP, similarity in bioactivity assessed from HTS/HCS data can still provide valuable evidence towards justifying the grouping, as discussed in Section 2.4.2. For endpoints that are less well understood or involve a complex AOP network, uncertainty in using mechanistic approaches for data-gap filling increases.

For endpoints where the mechanism is well understood, the use of a grouping approach may not be needed. For example, the DAs for skin sensitisation (described in OECD Guideline 497), may be a more practical means of data-gap filling. OECD Guideline 497 includes two Integrated Testing Strategies (ITSv1 and ITSv2) within the DA for skin sensitisation, which incorporate predictions from *in silico* tools, in combination with *in vitro* (human cell line activation test (h-CLAT)) and *in chemico* (Direct Peptide Reactivity Assay (DPRA)) predictions for hazard identification and classification based on potency prediction (i.e. GHS Cat 1A or 1B). The ITSv1 and ITSv2utilise predictions from Derek Nexus and the QSAR Toolbox, respectively. These *in silico* tools use structural alerts based on electrophilic binding to skin proteins for the parent substance and potential products following skin metabolism and auto-oxidation. In cases where no alert is identified following the ITSv1 DASS, Derek Nexus evaluates the reliability of the negative prediction by identifying any novel fragments in the chemical and/or any fragments which are commonly mispredicted by the structural alerts (Chilton et al., 2018). In cases where no alert is identified when following the ITSv2 DASS, the QSAR Toolbox attempts to identify potential analogues and fill the data gap through read-across as part of the DA. Both strategies are only appropriate when the target substance falls within the applicability domain of the individual information sources (h-CLAT, DPRA and *in silico* prediction; see Supporting Document to the OECD Guideline 497 on Defined Approaches for Skin Sensitisation for more detailed discussion of applicability domain; OECD 2023b).

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

- The identification of compounds as members of a category provides an insight into the potential effects of the compounds that might otherwise be overlooked.
- Reviewing data across members of a category may uncover inconsistencies e.g. a poorly documented study from one member could be strengthened by higher quality studies from other members; a study demonstrating different effects for one member may be an indication of a breakpoint in the category.
- The use of a category approach may also provide significant advantages in the evaluation of compounds that are often considered as “difficult”, in the sense that these can present technical difficulties when carrying out standard test protocols (examples are given in (Hart, 2007) and (Comber and Simpson, 2007)).
- In order to gain future efficiencies, category proposals may be expanded via the inclusion of chemicals that may be addressed under various global programmes.

Use of a category approach can also provide significant efficiencies and benefits when identifying and filling data gaps. A category test plan is designed to provide information to characterise the category as a whole, rather than to fill every data point for every chemical in the category. This reflects a more efficient approach than plans for obtaining data on individual chemicals. Knowledge of the expected biological effects of the category will be helpful in deciding if testing is needed and the nature and scope of the test to be carried out. One example of such a test plan is featured within the US EPA’s National PFAS Testing Strategy, where a large inventory of PFAS were subcategorised into smaller structural categories, and the

medoid of each was selected as a plausible candidate for further testing and evaluation (EPA, 2021[28]; www.epa.gov/assessing-and-managing-chemicals-under-tsca/national-pfas-testing-strategy; Patlewicz et al., 2024a). Where the goal is confirming an effect (e.g. skin irritation or corrosion) or the absence of particular property (e.g. acute oral toxicity) for an individual category member, a simple *in vitro* test might provide adequate confirmation. This approach is further elaborated on in Section 2.4.2 through the concept of bioactivity similarity using HTS/HCS data or alternatively using *in silico* in profiling approaches (see Section 2.4.4).

Another benefit of using a category approach is that it allows for an evaluation of the biological basis for the effects seen in a group of chemicals within a category. In certain cases, it may be feasible to elucidate common MOA, but at minimum, some evaluation of similarity in bioactivity can be made by profiling on the basis of HTS/HCS if such empirical data is indeed available. If it is known that members of a chemical category share a common MOA, the confidence in that category will be significantly greater than that associated with the use of an analogue approach where the MOA or the pairwise bioactivity similarity of the target and source chemical is unknown. This confidence increases with increasing numbers of chemicals and with empirical data included in the category.

In the vast majority of cases, a MOA/AOP will not be known, given the limited examples that have been published to date in platforms such as the AOPWiki (https://aopwiki.org/). Profiling members of a category could be approached based on their bioactivity similarity, reactivity similarity, metabolic similarity, etc., taking advantage of *in silico* tools (alert rule bases, or (Q)SAR model predictions) as well as generating specific *in vitro* or *in vivo* supporting data. For a large category[29], both the presence and absence of certain hazards, as well as the trend of an effect across a category, can be readily assessed. The consistency and concordance across members and between endpoints can provide a basis upon which the properties of individual members of the category can be identified with the necessary confidence.

For more limited comparisons, particularly with chemicals containing multiple functional groups, it may be harder to obtain the same level of confidence. For filling data gaps, a category approach can provide significant advantages compared with the analogue approach in that the category approach permits analysis of trends in properties and evaluation of consistency and concordance amongst the category members.

Within an analogue approach, confidence might also be derived when both target and source chemicals display a consistent pattern across many endpoints. For example, a comparable pattern in the types of effects for target and source substances would likely be observed for different endpoints, including acute and subchronic/chronic endpoints. A consistent pattern between MOA for certain endpoints may also be expected, e.g. acute toxicity in fish and daphnids. For endpoints such as skin sensitisation and mutagenicity where covalent binding of substances to cellular nucleophiles such as DNA and skin proteins respectively play an important role, a correlation in outcomes could be expected e.g. beta-propiolactone, a beta-lactone that can act through an acylating route is both a Ames mutagen and skin sensitiser. Patlewicz et al. (2010) and Mekenyan et al. (2010) compared the underlying chemical mechanisms for mutagenicity and skin sensitisation in an effort to evaluate the role mutagenicity information could play as a predictor of skin sensitisation potential.

Data gap filling techniques between chemical analogues have been extensively used, albeit on an *ad hoc* basis, e.g. within the OECD Cooperative Chemicals Assessment Programme, formerly the OECD High Production Volume (HPV) Chemicals Programme, the OECD IATA Case Studies Project[30], the OECD's

[28] https://www.epa.gov/system/files/documents/2021-10/pfas-natl-test-strategy.pdf

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

[30] See work on Integrated Approaches to Testing and Assessment (IATA) | OECD

WPMN, the EU Existing Chemicals Programme or for Classification and Labelling in the EU, the US EPA Voluntary High Production Volume (HPV) Challenge Program, and under Canada's Chemicals Management Plan (CMP).

An important consideration in revising this guidance is to move towards a more systematic approach to data gap filling that can provide a greater degree of transparency and reproducibility in the result. Read-across tools and approaches have continually evolved to facilitate integration of emerging data sources and supporting data for the formation of chemical categories to support the identification of groups of chemicals as risk assessment priorities (Rovida et al., 2020).

2.3. Selecting analogues/Creating chemical categories and setting boundaries

2.3.1. Selecting analogues

There are several ways of identifying potential analogue chemicals (source chemicals) with data with which the target chemical can be compared. In some cases, the choice of a source chemical may be driven by practical considerations, such as whether similar chemicals are produced for similar uses by the same company (or sector group of companies). In this case, no formal identification techniques might be required. Having said that, a more formal search strategy might identify additional analogues for comparison, thereby increasing the robustness of the subsequent data-gap filling. A formal search strategy can take one of two forms, either unsupervised or supervised:

- An unsupervised structural similarity approach involves a search that uses some constitutional representation of chemical structures (e.g. a 1-dimensional (1D) chemical fingerprint or some other molecular features) and a metric (such as the Tanimoto index, or cosine similarity) that measures the degree of similarity. This metric is used as a threshold to limit the total number of source analogues returned.
- A supervised approach is one where parameters relevant to the endpoints of concern are used to identify similar analogues. Various approaches and tools are described in detail in Chapter 4.

Evaluation of analogues is a critical step. The rationales described in Section 2.1.3 provide a useful starting point to characterise the underlying hypothesis that will be used in support of a category or analogue approach.

General considerations for evaluating analogues include an assessment of the physicochemical, reactivity, bioactivity, and metabolic similarity (ECETOC, 2012; Blackburn and Stuard, 2014; Schultz et al., 2015; Patlewicz et al., 2018). Software tools such as the QSAR Toolbox (OECD, 2009b) or Generalised Read-Across (GenRA) (Patlewicz and Shah, 2023) can be helpful to systematically compare analogues relative to these similarity contexts. The QSAR Toolbox contains a rich compendium of structural-alert-based profiling schemes, whereas GenRA contains structural and bioactivity profile representations that can be used to compare analogues or to set boundaries when creating new categories. These considerations as well as other software tools are discussed in more detail in Section 2.4.

2.3.2. Category and subcategory membership and applicability domain

Category membership and applicability domain

In an ideal situation, a category would identify all potential members when first developed. A top-down subcategorisation using clustering[31] techniques or substructural searching on a chemical inventory would

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

facilitate such an exploration. However, this situation is difficult to achieve in practice unless the boundaries and scope of the inventory are clearly defined. Practical constraints might include procurability and screenability of the substances, as well as the availability of relevant data.

Inclusion or exclusion of certain substances can introduce bias to the data gap filling; therefore, the constraints surrounding the source chemical inventory and the impact that this might have on the category members identified should be described to clarify assumptions. A clear category definition and description should allow a category to be expanded with additional substances.

The category definition includes the category name and the identity of category members (i.e. International Union of Pure and Applied Chemistry (IUPAC) chemical names, identifiers such as Chemical Abstracts Service (CAS) numbers, and structures). The category description should include a summary of common features: boundaries, physicochemical properties, if applicable allowed variations in chemical structure, and if known, any restrictions (e.g. variations that would change the effects of a substance significantly compared to the other substances in the category). The identity and purity of the test material, when experimental data are available, should be known. Ideally, the identity, constituents and content of impurities would also be known, especially if that might have a bearing on the toxicity profile or the classification and labelling of the category members.

Practical considerations will often influence the choice of chemicals included in the category. The selection of chemicals that are included in a particular chemical category is frequently guided by which chemicals are manufactured or imported by a registrant company or consortium company in the case of regulatory submissions. The successful use of a category approach should lead to the identification and characterisation (qualitative or quantitative) of the hazard(s) in question for all the members of the category, irrespective of their production volume or whether or not these are produced by the registrant companies carrying out the category evaluation. The practical considerations should not take precedence over the toxicological reasoning for grouping. Otherwise, this can introduce bias to the data gap filling.

There are significant potential advantages associated with the evaluation of a category that contains a high proportion of its likely members. Common factors that might influence the evaluation of a category include the degree of similarity between the members of a category and the quality/quantity of data for the members of that category. Indeed, if a large number of members have large data variability, false interpolation will be an issue. In the ideal case, the conclusions drawn from a category evaluation are likely to be more robust, since the evaluation is less prone to be affected by the subsequent addition of other chemicals, and the potential advantages of limiting animal and other testing are also likely to be greater.

A chemical can potentially belong to more than one category. For example, a multifunctional compound can belong to a category based on functional group A, as well as to another category based on functional group B. The properties of the compound will be influenced by the presence of both functional groups. Which category is more relevant will depend on the endpoint being considered and (Q)SAR approaches may be helpful to assign membership in such cases. This is also the case for UVCBs which are identified by their starting material and manufacturing process, and thus, a generic description of these parameters should be included in UVCB categories. Small changes to the same process can produce products with the same CAS number, but with differences in constituent composition. Therefore, depending on constituent composition, UVCBs with the same CAS number can belong in different categories, which also shows the importance of appropriate definition and description of the category and category members.

If a chemical is assessed and subsequently identified as a potential member of an existing category, it may be necessary to evaluate both the data for this chemical in light of the category evaluation and the category evaluation in light of the data for the additional chemical. If the initial category evaluation is sufficiently robust, the additional data are unlikely to alter the conclusions of the initial evaluation, but additional data may strengthen the category further. Since subsequent assessments of additional members of a category are possible at any time, there is an incentive to ensure that as many potential members of a category are included in the initial evaluation as possible. This would ensure that the evaluation is sufficiently robust in

order to minimise potential revisions resulting from adding data at a later date. Experience has shown that, in many cases, additional chemicals identified fall on either the lower or upper boundary[32] of an existing category. In those cases, additional testing or statistical analysis might be necessary to confirm that the chemicals belong to the category. In these cases, best professional judgment and WoE (See Chapter 3 Section 3.5) are used together in making recommendations/decisions about the level of testing that may be required, if any.

When assessing whether a chemical could be a new member of an existing category, clarity on the applicability domain of the category (i.e. which chemicals are covered by the category assessment) is important for the regulatory acceptance of the hazard conclusions. The applicability domain of a category would ideally identify the structural requirements and ranges of physicochemical, environmental fate, toxicological or ecotoxicological properties within which reliable estimations can be made for the category members. For this reason, the precise composition of the category (e.g. carbon number range, branching and position of branching, aromatic content, cyclicity, position and frequency of double bonds, functional group(s) of category members) should be defined where possible to set the boundaries that are used as inclusion/exclusion criteria. The applicability domain boundaries might also be supported by the commonality in MOA/AOP – some of which may be encoded in structural features characterising MIEs or through using a pragmatic threshold for bioactivity similarity. For example, there may be a trend of increasing acute aquatic toxicity with increasing chain length from C2 up to a carbon chain length of C12, after which no acute aquatic toxicity is seen because the water solubility has decreased with increasing chain length. Thus, it is critical to ensure the applicability domain for aquatic toxicity accounts for the change in solubility. There is a breakpoint in water solubility for many organic substances that must be incorporated in the applicability domain.

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

Subcategories

In some cases, an effect can be present or follow a trend for some but not all members of the category. An example is the glycol ethers, where methyl, ethyl, di-methyl and de-ethyl ethers show reproductive toxicity, but larger molecules are not (OECD, 2004b). For other properties/effect types, the category may show a consistent trend where the resulting potencies lead to different classifications. Examples include the lower aliphatic ethers, where aquatic toxicity is insufficient to lead to classification for aquatic toxicity with the lower members of the category but does lead to classification for this effect with higher members (Hart and Veith, 2007).

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

- An effect that varies in intensity across the category, such that some members of the category meet the criteria for one hazard classification for the particular endpoint, whereas other members of the category meet the criteria for another. These subcategory definitions can be:
    - Qualitative (i.e. these have degrees of hazard potential or different regulatory classifications).
    - Quantitative (i.e. the numerical values of the endpoint include values on either side of a breakpoint).

[32] Category members falling at the opposite extremes of a trend and within which interpolations are considered reliable are called sentinel or boundary chemicals (OECD, 2007).

- An effect where there is a peak in activity or a breakpoint in a trend can also lead to the formation of subcategories.
- It is possible that a trend analysis may apply to a subcategory but not to the whole category.

The concept of subcategories has been introduced to improve the practicality and flexibility of the category approach and it does not alter the scientific basis of this approach. The organisation of the category in subcategories should be presented and justified in the category justification.

Subcategorisation can be useful to further refine the degree of similarity of substances within a group where a greater degree of similarity might lead to an improvement in trends of chemical/biological activity. However, this approach is frequently a compromise between degree of similarity and the number of substances present in a chemical category. If the subcategorisation results in a small number of substances remaining, then the resulting lack of data will thwart subsequent analysis (e.g. read-across, trend analysis).

It should be noted that considering subcategories on account of breakpoints in endpoint trends is distinct and different from excluding a potential candidate member due to its similarity not being consistent or conserved across different contexts and endpoints.

Examples for subcategorisation that have been encountered within the OECD Cooperative Chemicals Assessment Programme include the case of mono-, di-, tri-, tetra-, and penta- ethylene glycols, when a subcategory was denoted by a cut-off of chain length of 6-8 to account for the change in physical form from liquid to solid and a decrease in uptake (OECD 2004b). A slightly different approach was used in the case of oxo alcohols C9 to C13 where clear trends in properties were seen with increasing chain length (Caley et al., 2007). For environmental hazards, two category members exhibited higher ecotoxicity than the other five members and thus formed a subcategory in the assessment. For the long chain alcohols (C6-22 primary aliphatic alcohols), decreasing water solubility and increasing lipophilicity was observed with increasing chain length, leading to a cut-off for acute aquatic toxicity effects at C13 to C14 and around C15 for chronic effects. At C>18, biodegradability was reduced (OECD, 2006b).

Categories for human health or for the environment

Sometimes the category approach may be applicable and justified for human health endpoints (e.g. same functional group, same metabolism, or same MOA), but not for environmental endpoints (e.g. different environmental fate, different aquatic toxicity across members of the category), and vice versa. An example includes the C2-C4 aliphatic thiols category where hazardous properties identified for human health are identical across the four members of the category (irritation, skin sensitisation, repeated-dose toxicity (OECD, 2010a)). Category members also share identical acute aquatic toxicity properties, but the environmental fate properties differ and result in different hazard conclusions. Another example could be a homologous series of alkanes; the aquatic toxicity would be expected to follow a trend based on chain length, but differences could be expected for human health effects because of metabolism. Hexane and pentane are examples. Hexane's toxicity is mediated by its metabolite hexane-2,5-dione, whereas pentane is hydroxylated to its corresponding alcohol (ECETOC, 2012). If no additional information were available, these would be "outliers" in an otherwise homologous series. The appearance of outliers brings extra uncertainty to the prediction and their exclusion should be accompanied with appropriate understanding[33].

[33] Outlier here refers to difference in metabolism between two members that results in a divergence in the underlying basis for the toxicity. In (Q)SAR terminology this could be akin to an activity cliff. In analogue/category terminology, such an outlier represents a breakpoint. Outliers can also be experimental outliers due to large variations in experimental data.

A category approach could also be applicable to many human health effects where metabolism plays a role, but not be applicable for local effect such as skin/eye irritation which are not expected to be dependent on parent chemical metabolism.

Overall, the coverage of a category should not be assumed to be appropriate for all endpoints, rather, an endpoint-specific reasoning and justification is necessary.

2.4. Evaluating analogue and chemical category approaches with other similarity context information

2.4.1. Principle and considerations

A category of chemicals with similar structures is expected to show either the presence or the absence of a particular effect across the members of the category, based on a common functional group, physicochemical properties, common reactivity, metabolism, common bioactivity profile and a presumed MOA (if available). Modulation of effects could appear as a result of a constant change in chemical structure or physicochemical properties across the category, e.g. increase in acute toxicity as a result of increasing *log* K_{ow} across the category members. Examples can be found from the Cooperative Chemicals Assessment Programme (e.g. C8-C12 aliphatic thiols category) [34].

When a category or analogue approach is being applied, the read-across should be substantiated for every endpoint for which there is a data gap. A chemical category approach could also apply to several toxicological endpoints, since the structural changes across the category might affect changes in physicochemical properties, molecular descriptors or profilers that in turn would result in changes for several toxicological properties (or other endpoints) of the individual category members. However, it may only be possible to identify the trends and changes for some, and not all, of the endpoints of potential interest. For example, it might not be possible to identify a trend in a case of too-low data density in the category or an inadequate distribution of available data across the category members, e.g. the data does not cover the boundaries of allowed structural variations within the category.

When the data for a category include one or more exceptions to the effects expected from a similar bioactivity profile, a review of the toxicological data for the category should generally be able to explain the difference in toxicity. The uncertainty of the grouping approach increases – possibly beyond an acceptable level for the considered purpose – if the exceptions cannot be explained. However, exceptions should not systematically be excluded from the category since the information or experimental data that provide can explain certain characteristics observed (e.g. absence of trend for a given endpoint) and may guide on the best approach to take for filling data gaps (e.g. worst case read-across versus read-across to the closest analogue). The presence of such “outlying” effects underlines the importance of developing an understanding of the MOA within categories. The use of negative and positive control substances is helpful in this context, which are chemicals with a relatively well-defined MOA that cause characteristic changes. The similarity (or dissimilarity) of effects between a negative or positive control chemical and a test substance can help to build confidence in a grouping hypothesis.

A category may be developed and/or justified on more than one basis. For example, a category could be justified by both chain length and metabolic pathway (Caley et al., 2007). In (Helman et al., 2018), analogues were identified by physicochemical information in conjunction with structural information to form categories whereas in (Tate et al., 2021), analogues were identified on the basis of targeted transcriptomics data in conjunction with structural information. In (Lizarraga et al., 2019), analogues were justified on the basis of ToxCast data to strengthen the biological plausibility in concert with empirical metabolism data,

[34] https://hpvchemicals.oecd.org/ui/Default.aspx

physicochemical data and structural information. Multiple arguments for justifications can increase confidence in the category, likewise, contrasting information or information on differences in biological effect may decrease confidence. In (Sperber et al., 2019), metabolomics data were used to substantiate a read-across for repeated dose and reproductive toxicity. This increased confidence is largely a result of the more detailed evidence that the common MOA has been identified. In an evaluation by (Roe et al., 2024) of the 2630 testing proposals submitted to ECHA, 304 read-across hypotheses were identified. Of these, structural considerations were the foundational element in almost all cases.

Comparable considerations also apply to the analogue approach, where, in addition to structural similarity and similar physicochemical properties between the source chemical(s) and the target chemical, criteria such as common functional group, biochemical processes and MOA, or environmental fate come into play for judging the suitability of (a) source chemical(s). It is helpful to consider if the method chosen for analogue identification relates to the endpoint of interest or whether a more unsupervised approach is used to identify source analogues irrespective of endpoint that will be filled through read-across. For example, the presence of substructures known to be associated with covalent protein binding mechanisms might be useful to identify analogues for skin sensitisation but might not be relevant for other toxicological endpoints. There are several structure-based chemical similarity methods, but all methods contain two key components: firstly, an approach to characterise or define the chemicals within the chemical space and secondly, a method of calculating the degree of similarity between two chemicals, typically a mathematical formula (e.g. Tanimoto). It is possible to use a subset of descriptors to define similarity. The subset could be determined by expert knowledge, machine learning (e.g. genetic algorithm), or by a combination of the two.

One example is the use of a metabolic pathway approach, where the category approach will be able to address the common toxicological mechanism for endpoints related to systemic effects related to the metabolite(s), whereas it may not predict the local effects (on skin and other membranes) caused by the parent compound. Another example is the category of monoethylene glycol ethers and their acetates or diethylene glycol ethers and their acetates (OECD, 2004b, Ball et al., 2014). Another example is alkaline properties driving the acute oral and dermal toxicity and therefore justifying the grouping of primary amines, whereas differences in metabolism (owing to structural differences) between members of the category (i.e. methylamine and tert-butylamine differ from the rest of the category) lead to different patterns of effects for chronic toxicity (see C1-C13 primary amines) (OECD, 2011b). In this example the metabolites are causing the observed toxicological effect and thus the formation pattern predicts the observed toxicity. In (Boyce et al., 2023), the metabolic similarity of a group of aromatic amines were considered taking into account the similarity in their transformation pathways as well as the similarity between the simulated metabolites.

For some series of compounds, the lower or upper end of the series may show marked changes in effects. For example, at the lower end of a series of alkyl compounds with varying chain length, the methyl analogue may have exceptional properties. For example, methyl alcohol and ethyl alcohol exhibit differences in their acute toxicity. Differences are seen in the carcinogenic profile between butter yellow and its ethyl homologue as well as between methylcarbamate and ethylcarbamate. This may be the result of specific differences in the route of metabolism (Jäckh, 2007). It is important to point out that several examples are needed to prove that an aberrant toxicity value is not due to the uncertainty associated with a specific data point.

The presence of a breakpoint (e.g. structural change, change in physicochemical properties) can indicate a change in the MOA, ADME properties, or the effect of a consistent trend across a category. In a homologous series of organic compounds, there is often a breakpoint, e.g. the loss of aquatic toxicity as carbon chain length increases and solubility decreases. The use of additional chemicals that can serve as positive and negative controls have the potential to contribute to the justification of chemical category and read-across.

The importance of a common MOA is also a factor in deciding what chemicals would not be expected to be members of a category. Variations in chemical structure can affect both toxicokinetics (e.g. uptake and bioavailability) and toxicodynamics (e.g. interactions with receptors and enzymes). For example, the introduction of a carboxylate or sulfate functional groups often decreases bioavailability and toxicity to mammals, while halogen substituents tend to increase lipophilicity and increase toxicological activity (see example by (Jäckh, 2007) in (Worth and Tier, 2007)). Thiols and esters are not considered as relevant analogues for evaluation of ether activity (see example in (Hart and Veith, 2007)). Variation in chain length within the same class of PFAS carboxylates has been shown to lead to differences in their toxicokinetic profiles (Smeltz et al., 2023). Examples of how analogues are identified and evaluated are available in Chapter 4.

2.4.2. Use of high content/high throughput/omics screening data to inform the development and justification of grouping chemicals

This section introduces a variety of approaches for how biological information can be used in forming and justifying chemical groups. In particular, the concept of bioactivity similarity is described as assessing similarity of the biological responses when exposed to test chemical using high content and/or high throughput screening technologies, often as part of a bridging study of both source and target substances.

The definition of a group typically starts with structural similarity and allowed structural differences and then continues with investigating other contexts of similarity, including bioactivity similarity, though this is not mandatory. Two overarching options are introduced for deriving and using bioactivity similarity as a line of evidence for a grouping hypothesis. First, whether the bioactivity similarity is associated with any mechanistic evidence (or not); and second, whether the bioactivity similarity is directly associated with the endpoint being read across (or not). All options could potentially contribute evidence to a grouping hypothesis, though with differing levels of relevance depending upon the regulatory context, as introduced in Table 1. Furthermore, the quadrants in Table 1 can form a roadmap for data that could potentially be generated to transition from lowest to highest evidence level for the endpoint being read across.

This includes the level of evidence provided and examples of the approaches used. More formal requirements are defined by the regional legislation and associated guidance for the grouping/read-across. Case study 2019-1(OECD, 2020g) focuses on establishing a proof-of-concept for the value added by NAMs including bioactivity data to support the bioactivity similarity of analogues in read-across. Existing toxicodynamic data, including NAMs, are used to evaluate the bioactivity across the category. The starting hypothesis for this category is based on their highly similar chemical structure, the target chemical is expected to have similar bioavailability and bioactivity as the analogue chemicals. The similar chemical structure and physicochemical properties will result in similar bioavailability, metabolism, and reactivity, leading to similar biological and functional effects. The available *in vivo* systemic toxicity data generally demonstrate the similar biological activity across the category. *In silico* data, *in vitro* NAM data investigating ADME and bioactivity, and toxicogenomics supports that category members are expected to have the same or similar biological activity and MOAs responsible for the observed effect.

Table 1. Options for introducing bioactivity similarity as a line of evidence in a grouping hypothesis

		Bioactivity similarity is associated with mechanistic evidence?	
		No	**Yes**
Bioactivity similarity is directly associated with the endpoint?	**No**	Lowest evidence level, e.g. measured by High Throughput Phenotypic Profiling (HTPP), or omics screening with no mechanistic interpretation	Medium evidence level, e.g. measured by ToxCast assays, or omics including a plausible toxicological interpretation, for example of a mode of action that is presumed to manifest in a relevant endpoint
	Yes	Medium evidence level, e.g. using administered equivalent doses (AEDs) derived from HTS/HCS assays	Highest evidence level, e.g. measurement of KE biomarkers in known AOP, or omics including a plausible toxicological interpretation linked directly to endpoint

Source: Based on (OECD, 2020g)

Bioactivity similarity can be demonstrated from various data sources. In the simplest case (Table 1, top left) – bioactivity similarity could be shown using biological activity data from an assay or suite of assays that may not be mapped to a specific endpoint. For example, high throughput phenotypic profiling (HTPP) assay data provides some measure of bioactivity information but cannot be directly mapped to a specific adverse outcome that would be measured in a regulatory guideline test. HTPP uses a combination of fluorescent probes to label a variety of organelles and measures a large number of phenotypic features (> 1000) at the single cell level to detect chemical induced changes in cell morphology. The profile of the phenotype features themselves provides a convenient representation to compare chemicals – in the same manner that a Tanimoto or dot product cosine metric is used to derive a pairwise similarity on the basis of chemical features. Omics technologies, when applied as a 'profiling' approach (where a *profile* comprises the set of all measured features or a subset of statistically pre-filtered features, (i.e. it is data-driven and does not use any external toxicological knowledge) and no mechanistic interpretation of the omics data is conducted, can also generate data for bioactivity similarity assessments in the top left quadrant of Table 1. This approach can be referred to as bioactivity *profile*-based grouping (Gruszczynska et al., 2024), in which metabolomics and transcriptomics profiles are used to identify an analogue for read-across. Beyond a single pairwise comparison, cluster analyses can also be performed to visualise the similarities across multiple substances to determine whether they may be grouped together.

The next case (Table 1, bottom left) continues with the example of HTS/HCS data streams in which chemicals are typically screened at multiple time points in a concentration response format to elicit an *in vitro* threshold for bioactivity, which can be converted to administered equivalent doses (AEDs) using reverse dosimetry. Such HTS/HCS derived AEDs could then be compared to effect levels from traditional *in vivo* studies. This would provide a means to quantify similarity on the basis of potency for an endpoint across source analogues relative to the target but without a direct mechanistic basis.

Greater confidence in grouping based on bioactivity similarity can be achieved by introducing mechanistic evidence (Table 1, top right). The Toxicity Forecaster program (ToxCast)[35] run by the US EPA represents a well-known example of using a suite of HTS assays to provide a biological profile for a chemical. Usually, no single assay is interpreted in isolation – rather assays are grouped by technology or vendor or gene to provide an overall perspective of likely activity. For example, assays that inform on estrogen receptor (ER) binding have been used to help rank and prioritise chemicals as part of the US EPA's Endocrine Disruptor Screening Program (EDSP). In most instances, the ToxCast suite of assays will align with the top right quadrant of Table 1 where some mechanistic information is more readily available, but the alignment to a

[35] https://www.epa.gov/comptox-tools/toxicity-forecasting-toxcast

regulatory endpoint is less clear, aside from specific cases such as for endocrine effects and the targeted set of assays that are grouped in a defined approach and aligned to an AOP to address the endpoint of concern. Omics assays could also inform a tiered approach of transitioning from the top left through to top right to bottom right quadrant. For instance, a plausible toxicological interpretation of the molecular responses relating to a MOA - which is presumed to manifest in an endpoint - can provide mechanistic evidence towards the grouping justification, placing the approach in the top right quadrant. For the case where the toxicological interpretation of the molecular mechanism is strongly associated with the endpoint being read across (for example, the molecular KE biomarkers measured by the omics assay are part of an established AOP leading to that endpoint), this qualifies for the lower right quadrant in Table 1. Of particular relevance here is where the omics input data take the form of a 'signature', which comprises a pre-specified, reduced (i.e. targeted) set of measured features that are associated with one (or more) MOAs, molecular pathways, AOPs, or endpoints, i.e. it is knowledge-driven using external toxicological knowledge. This approach is referred to as bioactivity signature-based grouping. Recently, predictive gene set signatures that align with events within an AOP, to anchor the observed transcriptional changes with a potential adverse effect, have been developed. Example signatures include those useful for predicting DNA damage inducing chemicals (Buick et al., 2015; Yauk et al., 2016) and ER α modulation (Ryan et al., 2016). Similarities in these gene expression signatures between a target chemical and source analogue can support the similarity justification.

A further strategy for increasing the confidence in the grouping justification (associated with the right-hand quadrants in Table 1) is using positive controls for a MOA or endpoint of interest, sometimes referred to as an 'anchor chemical' (see also Appendix, CG-ARM). Both HTPP and omics assays can be used to help infer a possible MOA for one or more chemicals based on a sufficiently high bioactivity similarity with a known anchor chemical such as a pesticide or drug that has a well-defined mode of action in terms of its toxicity and where the bioactivity is associated with that MOA. This scenario may be limited to those instances where membership to an existing category based on a common MOA was being evaluated. Here, a similar profile would provide some corroborating evidence that a potential new category member was indeed likely to act via a similar MOA.

The Connectivity Map (CMap) concept provides a further approach for supporting a MOA-based grouping hypothesis, based on a panel of anchor chemicals (De Abrew et al., 2016). The original proof of concept investigation with 4 cell lines and 34 well-studied chemicals was subsequently expanded to include 19 cell lines and 186 chemicals (De Abrew et al., 2019). This larger study identified four factors as having a significant impact on the grouping (promiscuity of chemical, dose, cell line and timepoint), which should be taken into consideration when interpreting the results from connectivity mapping. More recently, the CMap approach has been applied to demonstrate how a target chemical can be shown to exhibit a MOA that is dissimilar to a panel of anchor chemicals with defined endocrine disruptor MOAs (De Abrew et al., 2022).

The high content and/or high throughput data streams introduced here can be used in isolation to help substantiate the bioactivity similarity of members of a category or to provide a metric to compare source to target in an analogue approach. The data streams can also be integrated – information from HTPP or omics might inform on the phenotypes or molecular features to target in a more focused set of HTS assays – i.e. transitioning through the quadrants in Table 1 to increase scientific confidence and evidence. For example, ER effects observed in a transcriptomics study could be corroborated and confirmed in more targeted HTS assays that are linked to the associated AOP.

In addition, some of the potential mechanisms of actions elicited from these HTS/HCS data streams have been closely aligned to regulatory endpoints as part of an AOP (Table 1, bottom right quadrant). The assay battery supporting the assessment of endocrine activity under programmes such as the US EPA's EDSP provide a good example of a data stream with higher evidence and scientific confidence.

Beyond using the assays to substantiate and provide comparisons for specific source analogues identified through structure-based methods, a pre-existing compilation of profiles can also be used as a resource to

search for ‘similar’ analogues where similarity is defined by the feature profile itself. An example of how this has been implemented is within the GenRA tool (see Section 2.4.3) where source analogues can be identified on the basis of ToxCast profiles to return the most similar candidates on the basis of those ToxCast outcomes. Hybrid options to return the most similar analogues on the basis of bioactivity similarity in concert with other similarity contexts such as metabolism or structure are conceivable. At the time of writing, GenRA provides the capability to identify substances that are similar on the basis of structural fingerprints and bioactivity profiles (either based on the full complement of assays or restricted vendor specific profiles). Attention should be paid to evaluate whether analogues identified have been tested in the similar suite and number of assays.

Providing evidence of the reliability of these alternative approaches for grouping chemicals helps to build their international acceptance. Recently, the MetAbolomics ring-Trial for CHemical groupING (MATCHING), comprising six industrial, government and academic partners, evaluated the inter-laboratory reproducibility of applying metabolomics to group eight test substances, based on the bioactivity profiles measured in rat plasma (Viant et al., 2024b). Each partner applied their preferred LC-MS metabolomics workflows to acquire, process, quality-assess, statistically analyse and report their grouping results. All of the labs that acquired high-quality metabolomics data correctly identified the grouping into three categories, demonstrating the reliability of this approach, and providing evidence that some heterogeneity in the metabolomics methods is not detrimental to consistent grouping. This consortium proposed best practices in using metabolomics for bioactivity-profile based grouping (Viant et al., 2024b).

2.4.3. Use of Adverse Outcome Pathways (AOPs) to mechanistically inform the grouping of chemicals

Chemicals can be grouped according to their ability to trigger the same MIE or KE, although elucidation of the full pathway from MIE to adverse outcome is not necessary for building a chemical category. To develop an AOP, establishing causal links between the MIE or KE used to group chemicals and the apical endpoint for which the data gaps are to be filled is necessary. There are many rule-based ‘profilers’ /(Q)SARs which have been implemented for use in the QSAR Toolbox or other computational workflows such as within Toxtree, as KNIME workflows, or within other (Q)SAR repositories such as OPERA, Danish (Q)SAR Database, or VEGA (see Section 2.4.4). Many of these provide convenient means of grouping chemicals on the basis of their commonality in presumed MIE. For QSAR Toolbox profilers, structural features are typically used to define the applicability domain, and to a lesser extent, physicochemical properties (e.g. MW for ER binding). Arguably, the most mature profiles are those for DNA binding (Enoch and Cronin, 2010), protein binding (Enoch et al., 2011), respiratory sensitisation (Enoch et al., 2012), and Developmental and Reproductive Toxicity (DART) profilers (Wu et al., 2013).

Results from conventional *in vitro* assays and other HTS/HCS assays can be helpful in grouping chemicals, particularly if assays can be mapped to the MIE or KE of an AOP. For example, the genetically engineered yeast-based bioreporter system for ER binding-based gene activation can be used to screen chemicals for the potential to be reproductive toxicants. In this scenario, the MIE is ER-binding, the adverse outcome is reproductive toxicity, and the gene expression is an intermediate KE. Hence, the justification for grouping chemicals to fill data data gaps on reproductive toxicity could be strengthened by showing that all chemicals in the category have the same results in the ER-binding assay, in addition justification on structural similarity of category members. As databases of the results of mechanistic assays are developed, it will be possible to derive new prediction models including (Q)SARs to determine which chemicals are likely to be active, based on 2D or 3D structural information. Such models have been developed for ER and androgen receptor (AR) activation as part of two collaborative modelling projects (Collaborative Estrogen Receptor Activity Prediction Project (CERAPP), Mansouri et al., 2016; Collaborative Modeling Project for Androgen Receptor Activity (CoMPARA), Mansouri et al., 2020). ER and AR predictions for chemicals tested in the ToxCast program can be found on the US EPA’s CompTox Chemicals Dashboard. The ER and AR consensus models are described in several publications and can be accessed via the OPERA suite

(Mansouri et al., 2018). SARs for ER binding have also been implemented in other software tools, such as the QSAR Toolbox, to facilitate profiling of substances. Other tools include structural alerts for androgen and estrogen endpoints based on protein and enzyme perturbation (Derek Nexus) or predict the androgen and estrogen binding capacities based on the assumption that distances between electrophilic sites in the receptor determine the requirements for the binding mechanism (TIMES). Interpreting an adverse effect from MIEs and KEs should be in the context of a particular adverse outcome and a particular pathway. For example, ER-binding can be mechanistically linked to reproductive toxicity, but chemical binding to the aryl hydrocarbon receptor, another possible MIE, does not necessarily lead to reproductive toxicity, nor are all reproductive toxicants ER binders (Beischlag et al., 2008).

If including data from non-guideline test methods such as from HTS or HCS assays, descriptions of the methods or cited links to sources that summarise the methods should be provided. A template for the description is available in the "OECD Guidance Document for Describing Non-Guideline *In Vitro* Test Methods, Series on Testing and Assessment" (OECD, 2014a). Examples of description using the template can be found in Joint Research Centre (JRC) EURL ECVAM Database[36] service on alternative methods to animal experimentation (DB-ALM) and US EPA Toxicity ForeCaster (ToxCast™) Data[37].

2.4.4. Interdependence between (Q)SARs and categories and support for chemical grouping

The chemical category and (Q)SAR concepts are strongly connected given that the underlying basis of both is essentially the same, i.e. toxicity is a function of chemical structure. The differences mainly lie in the formality of how that relationship has been developed and packaged, where global mathematical QSAR models usually are based on analysis based on a much broader training set, thereby possibly picking up global and local trends based on more substances. This has been extensively described in the literature. References include (NAFTA, 2012), (ECETOC, 2012), (Patlewicz et al., 2017), and (Patlewicz et al., 2018). The concept of forming a chemical category and then using empirical data on a few category members to estimate the missing values for the untested members is in essence an internal (Q)SAR. The reason this concept is so compatible with (Q)SAR is that this broad description of the categories concept and the historical description of (Q)SAR are the same.

The categories concept creates a practical and powerful approach for describing the structural requirements of toxicity mechanisms. Chemicals may be grouped together initially using expert judgment, which is reflected by the chemicals included. Further evaluation may question the similarity of some chemicals based on their empirical data and demonstrate evidence of anomalous behaviour, or other information about the chemical attributes may suggest some chemicals fit more than one category. Applying (a) (Q)SAR model(s) may also be useful to help substantiate the category based on the manner in which mechanistic information has been encoded, i.e. in providing the mechanistic insight to support the interpretation of the experimental data. For example, through using the structural alert profilers that exist within the QSAR Toolbox[38] or elsewhere. Recent examples include structure-based MechoA profilers for toxicity by (Bauer et al., 2018; Sapounidou et al., 2021; Firman et al., 2022).

(Q)SARs can also play a useful role in addressing uncertainties for substances that could fall into more than one category. For example, if a substance possessed two heteroatom functional groups, a (Q)SAR model fit to a relevant dataset which had instances of both functional groups could provide insight into which category a substance should be placed in.

[36] https://data.jrc.ec.europa.eu/dataset/b7597ada-148d-4560-9079-ab0a5539cad3

[37] https://www.epa.gov/comptox-tools/toxicity-forecasting-toxcast

[38] https://qsartoolbox.org/features/profiling/

Errors within a chosen (Q)SAR model for a specific chemical may exceed the inaccuracy in the potency estimate of the (Q)SAR model. For example, in ecotoxicity studies, some phenols are polar narcotics, some are uncouplers, and others are electrophilic. (Q)SAR models for each mechanism have comparable uncertainty, but the potency of the latter mechanism can be orders of magnitude greater than polar narcotics. The use of a category approach can thus help to ensure that the (Q)SAR estimates are based on scientifically valid models by aiding correct selection of the model. Further guidance is available in OECD Guidance Documents No. 49 (OECD, 2004g), and No. 69 (OECD, 2014c) which provide helpful guides for validating (Q)SAR technology for a variety of applications. In addition, the OECD Guidance Document No. 386 (OECD, 2024a) provides OECD principles for the assessment of (Q)SAR predictions and results based on multiple predictions with checklists including the assessment elements and practical examples. Further information on the use of internal (Q)SARs to express trends in categories[39], and on the use of external (Q)SARs to provide additional support for trends, is given in Chapter 3, Sections 3.3 and 3.4, respectively.

2.5. Uncertainty analysis

2.5.1. Background

There are a number of publications that provide insights on how uncertainty can be characterised and evaluated. EFSA guidance documents on uncertainty analysis in scientific assessments (EFSA, 2018a; 2018b) are a good starting point. The OECD overview document on IATA concepts (OECD, 2020a) describes two main types of uncertainty. The first relates to the data and methodological quality, including relevance, reliability, and completeness of the data. The second type comprises uncertainties in the interpretation, extrapolation, and integration of available data, including knowledge about the phenomena of interest (e.g. AOPs, exposure pathways), and methodological choices made. Indeed, uncertainty needs to be taken into consideration at different levels, from individual data sources, data interpretation steps, to regulatory conclusions. The “WHO Guidance on Evaluating and Expressing Uncertainty in Hazard Assessment” (WHO, 2018) offers additional support when considering uncertainties in quantitative hazard assessments.

In the application of WoE, the “OECD Guiding Principles and Key Elements for Establishing a Weight of Evidence for Chemical Assessment” (OECD, 2019a) provides universal guiding principles that should be considered when developing or augmenting systematic approach for WoE evaluations of chemicals, and the ECHA WoE[40,41] template provides a structured format and background information. A number of different frameworks and templates have been created to facilitate the identification of the uncertainties and strategies to resolve them, either by the application of assessment factors or by generating additional supporting data (OECD, 2020a).

[39] Internal (Q)SARs refer to the trends established amongst category members. Within the Toolbox, this type of data gap filling is also referred to as “trend analysis”.

[40] ECHA Template for Weight of Evidence/Uncertainty in hazard assessment, https://echa.europa.eu/documents/10162/17169198/template_for_weight_of_evidence_en.docx/eb183c2e-c360-cbce-7a58-ad2d1270e5bd

[41] ECHA, WoE/Uncertainty in hazard assessment, background document and example, https://echa.europa.eu/documents/10162/17169198/wo_eu_uncertainty_background_en.docx/4f2b49ab-ade0-6ee3-e977-8abe00c21c23

2.5.2. Main considerations related to uncertainties in grouping and read-across

The uncertainty in a read-across analysis addresses two main aspects - the quality and relevance of the data for the underlying analogues/category members, as well as the similarity (e.g. mechanistic, structural, metabolic similarities) rationales. Both aspects should be characterised and described sufficiently so that the robustness of the analogue/category approach can be evaluated in the context of the prediction, the problem formulation, and any regulatory requirements that may apply.

Useful considerations for addressing the uncertainty in a read-across assessment include:

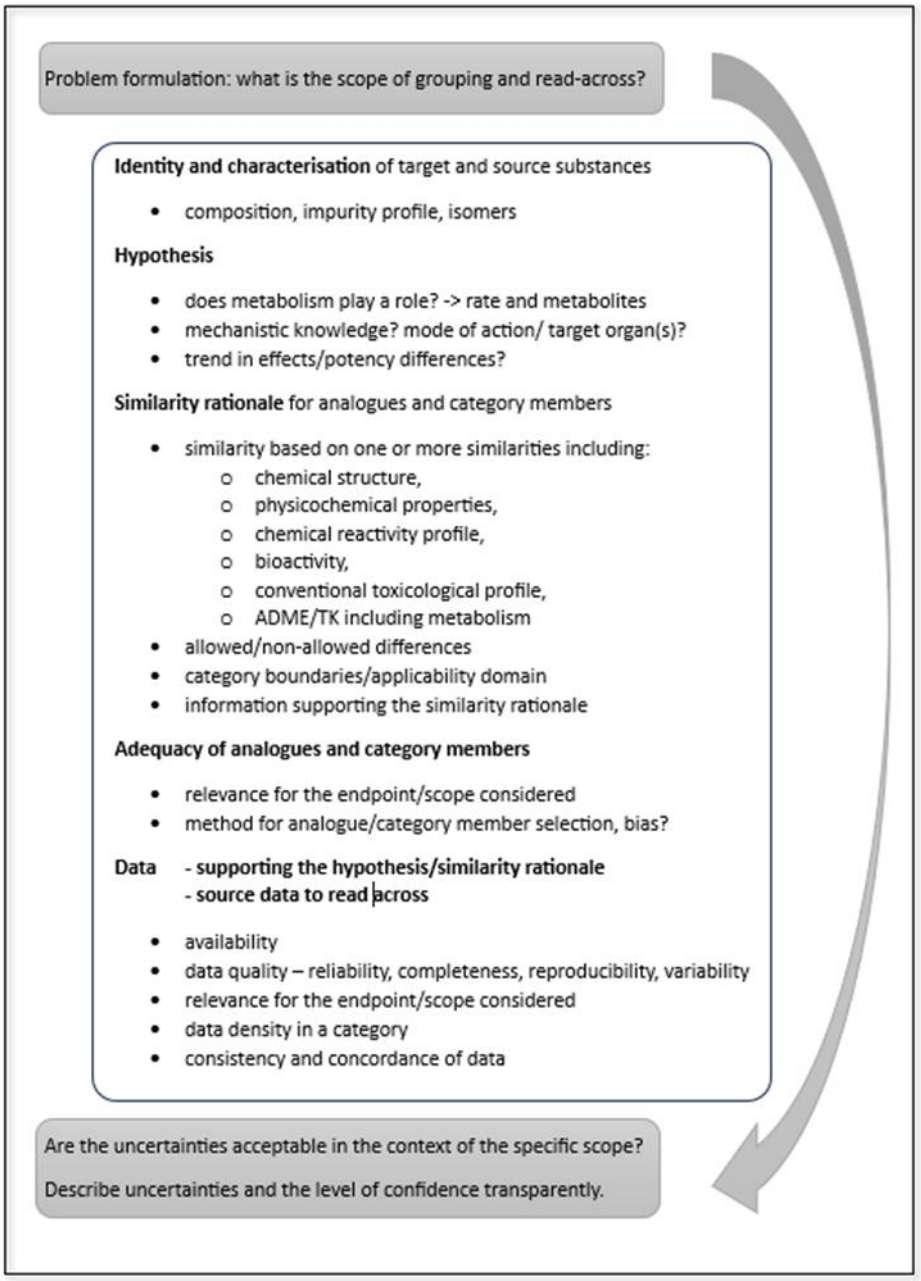
- Identification and characterisation (e.g. impurities) of source and target substances.
- Similarity in the structure and reactivity of analogue(s)/category members.
- Similarity or trend in the physicochemical properties.
- Similarity or trend in ADME properties, specifically the metabolic profile (metabolites, transformations, and sequence of).
- Similarity or trend in the biological activities.
- Applicability domain/boundaries of the category.
- Consistency and concordance in experimental and profiling outcomes among analogues.
- Consistency and concordance in experimental and profiling outcomes across endpoints.
- Number of analogues/the density and distribution of the category (both in terms of the chemicals represented and the data available across the category).
- Quality and relevance of the underlying experimental data for each of the endpoints covered.
- Presumed mechanistic basis underpinning the category or the analogue approach for a particular endpoint if available.
- Quality of the predicted properties generated by *in silico* approaches, if applicable.

There are several examples where templates have been created to help systematically identify and evaluate and document sources of uncertainty in read-across. Examples include the assessment framework proposed by Blackburn and Stuard, (2014) which followed an analogue identification and evaluation workflow developed by Wu et al., (2010). Cronin et al., (2022) developed a framework for characterising the uncertainty of structural alerts. Schultz et al., (2015) and Cronin et al., (2022) developed a set of reporting formats to facilitate an overall qualitative assessment of the read-across uncertainty. Patlewicz et al., (2015) created a template to capture sources and practical approaches to address uncertainty, informed by industry experience in developing read-across justifications for regulatory purposes, and closely aligned with technical guidance from the ECETOC Task Force (ECETOC, 2012). ECHA developed its own Read-Across Assessment Framework (RAAF) to provide a framework and guidance for consistent evaluation (ECHA, 2017b). Included in the ECHA RAAF, six scenarios are described, each comprising different assessment elements which in turn address different scientific considerations deemed critical to evaluate the validity and reliability of read-across. A review of several of these frameworks was described by Patlewicz et al., (2018) in an effort to compare and contrast their similarities and differences, as well as propose a harmonised scheme (Patlewicz et al., 2018). The intent was to demonstrate opportunities for quantifying uncertainties in different steps of the read-across workflow, as well as highlight where supporting information might reduce residual uncertainties or where alternative strategies for data gap filling might be most impactful. Schultz et al., (2019) also summarised the types of uncertainties encountered in read-across. There is a high degree of similarity between these frameworks in terms of the types of information that they aim to capture. Clearly, the RAAF is intended to evaluate EU REACH specific analogue and category use cases, whereas the other frameworks are more general, capturing various regulatory purposes such as hazard assessment or product stewardship applications. Their commonality predominantly lies in three main aspects – data quality, similarity contexts and grading scales. The most common of these frameworks are summarised in Section 2.6.1.

Overall, all frameworks highlight similar points for assessing uncertainties related to grouping and read-across approaches. A checklist of these main points is illustrated in Figure 3.

There are also several examples where templates have been created to facilitate the reporting of uncertainty within read-across. The most common of these are summarised in, Section 2.6.2.

Figure 3. Main aspects to consider for uncertainty assessment



The assessment of the quality of analogue/category member data and the similarity contexts underpinning the read-across is complex. Data quality is intended to capture the number of source analogues, robustness of the underlying data available for those source analogues and target, as well as the outcomes reported. The robustness of the underlying data covers not only the endpoint data that may be potentially

read across but all aspects of information available for the target and source analogues – starting from the validity of the substance identities. An understanding of the achievable accuracy of the available technical methods used to generate the data is also a consideration. Understanding the impact of data uncertainty, variability, and reproducibility is essential in evaluating the robustness of the underlying data.

The similarity contexts make reference to the different considerations that come into play in rationalising the relevance and validity of source analogues. Often structural similarity is the starting point in identifying source analogues/category members but evaluating the suitability of those analogues usually factors in considerations such as the similarity in reactivity profile, metabolic profile, physicochemical profile, bioactivity similarity, and common MOA.

For each of these similarity contexts, the evidence (*in silico* or empirical) available to substantiate similarity should be described, as well as the robustness of the supporting data. For example, for metabolic similarity, supporting evidence could include commonality in the metabolites formed, the types of pathways involved, and the rates of transformation. In the case of reactivity similarity, commonalities in the structural alert profile for features indicative of an electrophilic reaction mechanism or kinetic data to demonstrate similarity in rates of reaction could be described. If the data is predicted from *in silico* models, the concordance of those tools relative to the (Q)SAR models and prediction validation principles as described earlier should be considered.

The next step is to consider this in relation to the read-across being undertaken. Here an assessment of the level of concordance in effects and potency for a given endpoint across the analogues would be a consideration. Commonality in the effects noted and the severity of those effects across source analogues would also increase the robustness of any read-across performed. Good quality data for the source analogues would also increase the robustness versus data from poorly reported studies or those lacking key pieces of study information. Usually, the data matrix from source analogues will be incomplete hence a consideration of whether there is sufficient information from the available empirical and predicted data will also be a factor. Finally, since the development of an analogue/category approach often starts with an overarching hypothesis that defines how those analogues are identified, consideration of how the information gathered satisfies that requirement needs to be assessed. For these considerations the same scoring scheme can be applied to define an overall uncertainty for the read-across being proposed.

Whilst these considerations provide a means of identifying the sources of uncertainty and scoring them accordingly, a minimum or ideal overall score is beyond the scope of the guidance here as this will be dependent on the specific use case and regulatory purpose for which the read-across is being considered. For an extensive risk assessment case, there may be a minimum uncertainty that will be tolerated and if the uncertainty was too high, strategies to reduce or mitigate the uncertainties could be considered. These could take the form of assessment factors or alternatively new data could be generated to strengthen the component driving the overall score. For example, if the strength of evidence or the data quality for the metabolic similarity was the weakest aspect, additional supporting data might be generated or additional *in silico* predictions could be derived to strengthen that specific line of evidence.

Table 2 outlines the key components to consider when characterising uncertainty related to the similarity rationale. Within Table 2, evidence underpinning each similarity context is assessed on the basis of two aspects – the strength of the association between the similarity and the endpoint being read across as well as the quality of the data characterising the similarity itself. The assessment is performed using a qualitative grading scale ranging from low to high or insufficient. An overall score is then defined to weight the total uncertainty across all the similarity contexts for the source analogue(s).

An example of how this could be populated in practice is illustrated using the IATA Case Study for 90-day rat oral repeated-dose toxicity of chlorobenzene-related chemicals (OECD, 2020h).

The main aspects to consider overall for assessing uncertainties related to a grouping and read-across approach are summarised in Figure 3.

Table 2. Key components of uncertainty related to the similarity rationale

Similarity rationale	Data quality	Strength of Evidence	Comments
Structural similarity	High	Low	A benzene scaffold with 1-6 chlorines and with the potential to be oxidised to epoxide form. Pairwise similarities ranged from 0.58-0.92 amongst the members
Physicochemical property similarity	Medium	Low	The number and position of chlorine atoms affected the *log* K_{ow} which decreased with increasing number of chlorine substituents
ADME/TK including Metabolic similarity	Medium	Medium	Based on *in silico*, *in vivo*, and *in vitro* data, all members are expected to undergo similar metabolism, namely, epoxide hydrolysis and GSH conjugation. However, it was difficult to find experimental data regarding the detection of metabolites such as epoxide and quinone for all members. Furthermore, hexachlorobenzene did not show active metabolites in metabolite prediction. In terms of adverse effects of each metabolite, it was difficult to explain the toxicological effects of all predicted metabolites
Bioactivity similarity	Medium	Medium	Members mainly showed similar toxicological findings such as increased relative organ weight, hypertrophy, and histological changes in the liver and kidney.
Chemical reactivity profile			NA
Mode/Mechanism of Action similarity	Medium	Medium	Chlorobenzene-related chemicals induce hepatotoxicity and renal toxicity after repeated oral dosing in rat. Metabolites cause oxidative stress, GSH downregulation, protein binding, leading to toxicological effects. Thus, the chemicals that have the same metabolic pathways are expected to show similar key events. Furthermore, NOAEL of some chlorinated benzenes correlated with water solubility, because of the accessibility to P450 and cell permeability (molecular initiation events).
Conventional toxicological profile			NA
Overall uncertainty	Medium	Medium	All chemicals are expected to have similar active metabolites which lead to a key event. The existing data could not cover all the endpoints with some exceptions, such as target organs, in repeated dose toxicity test.

Note: NA – not applicable
Source: (OECD, 2020h)

Quantitative assessment of read-across uncertainties

An alternative means of assessing the uncertainties in a read-across could instead form a part of a systematic data-driven approach. Here, read-across performance could be assessed quantitatively relying on standard cross-validation and external validation techniques using statistical performance metrics that are typically applied when evaluating (Q)SAR models. One such data-driven approach was derived by (Shah et al., 2016) who described a mean of performing a data-driven assessment of *in vivo* toxicity read-across using chemical fingerprints, bioactivity fingerprints, or both. The intent was to establish a baseline in performance measure for read-across performed using structural characteristics which would allow any increase in performance to be compared when incorporating additional contexts of similarity, i.e. was there a significant improvement in performance when physicochemical or metabolic information in conjunction with structural information formed the basis of the read-across. The read-across prediction was defined as the similarity weighted activity of the source analogues – akin to a many-to-one read-across but where the

contribution from each source analogues was weighted by its pairwise similarity to the target substance. The approach has since been implemented in the web application GenRA (Patlewicz and Shah, 2023).

Other means of investigating and characterising confidence have been undertaken by (Yang et al., 2021) who quantified confidence intervals around repeated dose toxicity points of departure (PODs) for analogue pairs and the impact this had on read-across prediction. Quantifying the variability of repeated dose toxicology studies is an active area of research (Pham et al., 2020) to make explicit the uncertainty around point estimates such as lowest observed adverse effect level (LOAEL) and no observed adverse effect level (NOAEL) values. The problem with the quality and variation of experimental data has been discussed recently by Chapkanov et. al. for two endpoints - skin sensitisation (Chapkanov et al., 2023) and repeated dose toxicity (Chapkanov et al., 2024)-. The authors analysed the data variation for the aforementioned endpoints and proposed a classification that took into account the uncertainty of the experimental data. (Escher et al., 2019) outlined different levels of uncertainties including NAM data where test readiness and validation status were considerations. Scores have been developed for test readiness which allows the accuracies and predictions of tests to be quantified (Bal-Price et al., 2018). The Dempster Shafer theory (DST) (Dempster, 1967; Shafer, 1976; Rathman et al., 2018), a Bayesian based decision theory approach, allows for a fully quantitative combination of various types of test data taking into account individual test performances in order to derive likelihoods of test data being correct. DST based algorithms can provide probabilistic estimates based on the combined quality and reliability of NAM data.

2.6. Chapter 2 Annex - Examples

2.6.1. Overview of uncertainty assessment in existing read-across frameworks

Table 3. Aspects related to uncertainties covered by different frameworks

Framework	Blackburn and Stuard (2014)	Schultz et al. (2015)	Patlewicz et al. (2015)	Schultz et al. (2019)	RAAF (ECHA 2017b)
Read-across hypothesis	Related to similarity and considerations on data as below.	Uncertainties associated with the mechanistic relevance, completeness and application of the read-across approach. Is the category relevant for the endpoint considered? Overall weight-of-evidence supporting the prediction.	Aspects that impact the similarity rationale underpinning the read-across hypothesis	Uncertainty related to the argumentation of the read-across; Approach taken Mechanistic plausibility Supporting argumentation Robustness of the argument from hypothesis to application of read-across.	What is the toxicant – formation ((bio) transformation) of common compound, different compounds with same effect? Mechanism of toxic action – qualitative and quantitative aspects. Mode of action hypothesis. Biological target?
Similarity	Close structural similarity of the analogues to the target. Consider differences to target chemical: structural characteristics, physicochemical properties, metabolism.	Chemical, toxicokinetic and toxicodynamic similarity. Description of similarities and differences (structure, physicochemical properties, toxicokinetics, transformation, metabolic pathway, activation to reactive chemical species, bioavailability, biological and toxicological properties, similar toxicophore) and link to the read-across hypothesis.	Differences in structure and physicochemical properties influencing toxicity. Physicochemical, metabolic transformation and reactivity considerations related to similarity rationale.	Justification of similarity between target and source compounds: similarity in chemical structure, physicochemical properties, toxicodynamics, toxicokinetics (ADME); consider relevance of dissimilarities.	Similarities and differences (structure, physicochemical properties, fate e.g. degradation, bioaccumulation potential), link with the predicted property. Allowed and non-allowed differences, impact of impurities.

Framework	Blackburn and Stuard (2014)	Schultz et al. (2015)	Patlewicz et al. (2015)	Schultz et al. (2019)	RAAF (ECHA 2017b)
	Consider bridging data between analogue and target substances.				
Supporting data: Data quality Data adequacy	Sufficient quantity and quality of highly concordant data, (patterns of toxicity, range of potency).	Assessment of data associated with the similarity. Data matrix of key properties (structure, chemical properties, toxicokinetics, transformation/metabolism, biological and toxicological properties) of the analogues.	Uncertainties associated with the underlying data supporting the analogues. Sufficient quality and detail of description. Concordance of effects and potency across dataset and analogues.	Strength or robustness of the supporting data sets, including performance of the methods (reliability, accuracy, precision, repeatability and reproducibility).	Assessment of the adequacy and robustness of the scientific reasoning and of the supporting evidence. Consistency of effects in the data matrix. Bridging studies.
Source data: Data quality Data adequacy	Sufficient quantity and quality of highly concordant data, (patterns of toxicity, range of potency).	Appropriate toxicological studies of sufficient data quality to allow a meaningful read-across?	Uncertainties associated with the underlying data of the source analogues.	Uncertainty related to the data for the endpoint under consideration for the source compound(s). Data reliability, accuracy, precision, repeatability and reproducibility.	Reliability and adequacy of the source study/ies. Bias e.g. in selecting source study/ies.
Considerations on regulatory context/ level of acceptable uncertainties	-	Problem and premise of the read-across, overarching scenario. Is the uncertainty acceptable for a specific regulatory purpose?	-	Uncertainty related to the regulatory use, the impact this will have on acceptable levels of uncertainty	Applies to read-across in the context of EU REACH.
Scores: Qualitative	Qualitative (low, low to moderate, moderate, high).	Qualitative (low, moderate, high).	-	-	"Assessment Options" to be assigned to each Assessment Element, i.e. scores 5-1

Framework	Blackburn and Stuard (2014)	Schultz et al. (2015)	Patlewicz et al. (2015)	Schultz et al. (2019)	RAAF (ECHA 2017b)
Quantitative	And quantitative uncertainty factors proposed (1, 3, 10).				(acceptable with high/ medium/ just sufficient confidence, not acceptable in its current form, not acceptable).
Template	Yes, questionnaire (see Example 1, Section 1.1).	Yes, tables for assessing uncertainty associated with similarity justification and associated with the overall read-across approach (mechanistic relevance and completeness for the read-across; see Example 2, Section 1.1).	List of scientific confidence considerations of a read-across justification.	List of questions to guide through assessing uncertainties (see Example 4, Section 1.1).	List of "Assessment Elements". i.e. scientific considerations for systematic evaluation (see Example 3, Section 1.1).
Comments	Questionnaire applies only to quantitative endpoints (such as with estimated NOAEL).	-	Informed by industry experience in developing read-across justifications for regulatory purposes and closely aligned with the ECETOC technical guidance.	-	6 predefined scenarios (2 for analogue approach, 4 for category approach) with adapted Assessment Elements.

2.6.2. Examples of templates/systematic approaches to assess uncertainties

The following examples are taken from the template used for the IATA case studies by the OECD IATA Case Studies project. They can support a systematic assessment of uncertainties related to the grouping and read-across approach by leading through the different aspects to consider.

Example 1: Reporting template of uncertainty

The example was developed by the experience of the OECD IATA Case Study Project based on (Wu et al., 2010), and (Blackburn and Stuard, 2014).

Type 1: Analogue suitability rating for read-across

Evaluation criteria [a]	Question[b]	Uncertainty[c]
Structure and reactivity	Do the target and analogue have similar structural features and chemical reactivity?	
Metabolism	Do the target and analogue have similar metabolic pathways?	
Physicochemical properties	Do the target and analogue have similar physicochemical properties?	
…….		
Overall "suitability rating"[d]		

a- Criteria used for evaluating the suitability of analogues
b- Question and answer used for evaluating the criteria
c- Description of the uncertainties in the answer to the question
d- Rank (suitable, suitable with interpretation, not suitable, suitable with preconditions) derived from the decision tree
Source: This table is based on the decision tree of the framework by (Wu et al., 2010)

Type 2: Uncertainty associated with the prediction of hazard using read-across

Analogue data set characteristics[e]	Question[f]
Number of analogues contributing data	
Robustness of analogue data set	
Concordance of effect(s)	
…….	
Overall uncertainty of read-across prediction[g]	

e- Analogue data set characteristics used for evaluating overall uncertainty of read across prediction.
f- Description of the evaluation results of the analogue data set characteristics obtained by answering the questionnaire of the framework.
g- Rank of overall uncertainty of read across prediction derived from the evaluation results of analogue data set characteristics (low, low to moderate, moderate, high) with the description of the reason.
Source: This table is based on the framework by (Blackburn and Stuard, 2014)

Example 2: Reporting template of uncertainty

The example was developed by the experience of the OECD IATA Case Study Project based on (Schultz et al., 2015).

Type 1: Parameters and associated uncertainty used to justify category membership

Justification parameter[a]	Data uncertainty[b]	Strength of evidence[c]	Comment[d]
Structural similarity			
Physicochemical properties			
Metabolic similarity			
Mechanistic similarity			
Trends in effects			
…….			

Overall uncertainty in similarity of category members[e]

a- Similarity parameter used for justifying the category.
b- Rank of uncertainty (low, medium, high) associated with underlying data used for analysis
c- Rank of consistency (low, medium, high) within the data
d- Description of the reason for the assignment of the ranks of the uncertainty and strength of evidence
e- Rank of overall uncertainty (low, medium, high) and description of the reason
Source: This table is based on the framework by (Schultz et al., 2015)

Type 2: Uncertainty associated with the prediction of hazard and dose-response using read-across

Factor[f]	Uncertainty[g]	Comment[h]
Number of analogues contributing data		
Robustness of analogue data		
Concordance of effects		
Concordance of potency		
Severity of effect		
…….		
Overall uncertainty of read-across (low, medium, high)[i]		

f- Uncertainty factor associated with the prediction of hazard and dose-response using read-across.
g- Rank of uncertainty (low, medium, high)
h- Description of the reason for the assignment of the ranks of the uncertainty
i- Rank of overall uncertainty (low, medium, high) and description of the reason
Source: This table is based on the framework by (Schultz et al., 2015)

Example 3: Assessing uncertainty following the ECHA Read-Across Assessment Framework (RAAF)

The examples are from the ECHA RAAF read-across scenarios 1 and 2 for analogue read-across. For a detailed description of the Assessment Elements (AEs) and other scenarios, see ECHA RAAF (ECHA, 2017b).

Uncertainties of the read-across approach can be assessed and described by systematically addressing the topics as listed by the AEs, which describe crucial aspects to be taken into account for an adequate read-across approach.

- AEs for Scenario 1 (analogue approach for read-across based on hypothesis for (bio)transformation to common compound (s))
  - AE A.1 Common AE: Identity and Characterisation of the source substance
  - AE A.2 Common AE: Link of structural similarities and differences with the proposed prediction
  - AE A.3 Common AE: Reliability and adequacy of the source study
  - AE 1.1 Scenario-specific AE: Formation of common (identical) compound(s)
  - AE 1.2 Scenario-specific AE: The biological targets for the common compound(s)
  - AE 1.3 Scenario-specific AE: Exposure of the biological target(s) to the common compound(s)
  - AE 1.4 Scenario-specific AE: The impact of parent compounds
  - AE 1.5 Scenario-specific AE: Formation and impact of non-common compounds
  - AE A.4 Common AE: Bias that influences the prediction

- AEs for Scenario 2 (analogue approach for read-across based on hypothesis that different compounds have the same type of effects)
  - AE A.1 Common AE: Identity and Characterisation of the source substance
  - AE A.2 Common AE: Link of structural similarities and differences with the proposed prediction
  - AE A.3 Common AE: Reliability and adequacy of the source study
  - AE 2.1 Scenario-specific AE: Compounds the test organism is exposed to
  - AE 2.2 Scenario-specific AE: Common underlying mechanism, qualitative aspects
  - AE 2.3 Scenario-specific AE: Common underlying mechanism, quantitative aspects
  - AE 2.4 Scenario-specific AE: Exposure to other compounds than to those linked to the prediction
  - AE 2.5 Scenario-specific AE: Occurrence of other effects than covered by the hypothesis and justification
  - AE A.4 Common AE: Bias that influences the prediction

Example 4: Assessing uncertainty following a list of questions

In (Schultz et al., 2019), 30 questions were formulated according to 12 types of uncertainty that may be associated with performing a read-across. They can be used to guide through an assessment of the uncertainties of a given read-across. The extent of the answers required to some, or all questions depend on the overall question being addressed.

Uncertainty in read-across	Questions to evaluate the uncertainty
The context of, and relevance to, the regulatory use of the read-across prediction as defined by appropriate problem formulation	• Is the regulatory purpose of the read-across prediction clearly defined? • Is the acceptable level or degree of uncertainty for the stated purpose defined? • Is the stated acceptable level or degree of uncertainty appropriate for the stated regulatory purpose?
Type of category/group including the definition of the applicability domain	• Is the read-across approach (e.g. analogue or category) clearly reported? • Are the target and source chemicals clearly identified? • Is the applicability domain of the analogue or category defined? • Do target and source chemicals fit within the defined applicability domain?
The premise or hypothesis of the read-across.	• Is the hypothesis on which the read-across is based clearly stated and presented in sufficient detail to be assessed?
Mechanistic plausibility including completeness of the understanding of the MoA or AOP	• How clearly does the hypothesis state the chemical and biological mechanisms underpinning the toxic effect being read across?

Uncertainty in read-across	Questions to evaluate the uncertainty
	• Is there sufficient experimental information provided to support the proposed chemical and toxicological mechanisms? • How extensively does the experimental information provided support the mechanistic plausibility and / or the AOP or MOA on which the read-across is based?
Similarity in chemistry	• Are the chemical structures (i.e. 2D structure, isomers, SMILES and molecular formula) reported for the derivatives used in the read-across? • Are the dissimilarities in chemical structure reported and are they toxicologically relevant? • Are the relevant molecular and physicochemical properties (e.g. for molecular size, hydrophobicity, solubility, volatility, degradation etc.) reported for the derivatives used in the read-across? • Are the dissimilarities in molecular and physicochemical properties reported and are they toxicologically (or pharmacokinetically) relevant?
Toxicodynamic similarity	• Is there sufficient and consistent toxicodynamic information provided to establish similarity in the hazard of the derivatives used in the read-across?
Toxicokinetic similarity	• Is there sufficient ADME information provided to establish toxicokinetic similarity for the derivatives used in the read-across? • Are any dissimilarities in ADME properties (and, as appropriate, metabolism / degradation) toxicologically relevant?
The quality of the apical endpoint data used to fill the data gap	• Is the performance (e.g. reliability, accuracy, precision, repeatability and reproducibility) of the data read across reported clearly? • Has the quality of the data to be read across been assessed and are they sufficient to meet the purpose of the exercise i.e. complete and of sufficient quality?
The consistency in the effects and severity of the apical *in vivo* hazard and their concordance with regards to the intermediate and apical effects and potency data	• Is the qualitative expression of the data reported and is it consistent among the source chemicals? • Is the potency of the hazard reported and is it consistent among the source chemicals? • What are the temporal relationships between relevant endpoints? • What are the dose–response relationships between relevant endpoints?

Uncertainty in read-across	Questions to evaluate the uncertainty
Strength or robustness of the supporting datasets	• How extensively are the relevant or key events either empirically measured and/or modelled by appropriate *in silico*, *in chemico* and *in vitro* data? • Is the performance (e.g. reliability, accuracy, precision, repeatability and reproducibility) of the supporting methods adequately reported?
The WoE supporting the prediction	• Is there consistency in the supportive information (e.g. structural alerts) between analogues or within the category? • How many and how large are the dissimilarities in the supporting information (i.e. data gaps)?
Documentation and written evidence provided	• Is the read-across prediction adequately documented? • Does the evidence support the hypothesis that the uncertainty is acceptable for the stated purpose (as per Question 1)?

3 Techniques or Methods for Data Gap Filling

Box 3.1. Chapter 3 summary

This chapter:

- Explains the main approaches that are used for data gap filling: read-across (including for continuous endpoints and trend analysis) and (Q)SARs.
- Provides details of computational tools that can facilitate the grouping procedures that are outlined in Chapters 4 and 5.

3.1. Introduction

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

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

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

- Read-across (Section 3.2). Section **Error! Reference source not found.** discusses how to develop read-across justifications for each endpoint and considerations for the validity of read-across.
- Trend analysis and use of computational methods based on internal models (Section 3.3)
- Use of computational methods based on external models (Section 3.4). This includes conventional (Q)SAR models or prediction models derived from *in vitro* data.

In principle, the above-listed non-testing techniques can be used to indicate either the presence or the absence of an effect or an estimated value (e.g. a relevant toxicity value such as a LOAEL or a NOAEL) for an analogue or a group of substances. However, this is highly dependent on the substance under consideration, the endpoint, the level of information already available, the regulatory purpose, and the confidence that can be derived from its interpretation. Consequently, the generation of additional

experimental data by strategic testing may still be required to inform the properties of category members and develop confidence in the approach considering the WoE of all of the information available. Such testing may involve the generation of supporting data to substantiate bioactivity similarity. The OECD Guidance Document No. 311 (OECD, 2019a) provides universal guiding principles that should be considered when developing or augmenting systematic approaches to WoE for chemical evaluation and key elements to formulating a systematic approach to WoE.

The use of these techniques is described in more detail below. None of them are typically used in isolation. Usually, these techniques rely on building a case with varying degrees of applicability in the context of both the analogue approach and the wider category approach. Experience from current practice shows that qualitative or quantitative read-across is already widely used and is a viable approach for regulatory purposes on a case-by-case basis. While computational approaches based on structure-activity relationships (SARs), quantitative SARs (QSARs), quantitative activity-activity relationships (QAARs) or expert systems[42] can also provide a basis for filling data gaps, past experience shows that additional supporting evidence is still often required for acceptance of these estimates. In line with the IATA concept, use of many or all of the techniques in an integrated fashion with inclusion of available experimental information may lead to the most robust result.

3.2. Read-across

The use of read-across is widespread across regulatory jurisdictions, particularly as a mean to fill data gaps for information requirements under specific regulations. The term “read-across” is a generic and much used phrase. However, all the examples of categories and analogue approach such as from the OECD HPV programme[43], the OECD IATA Case Studies Project[44] and regulatory applications within Member Countries, make it clear that read-across can only be used on a case-by-case basis by providing a hypothesis on which the read-across is based. Adequate justification, documentation (see Section **Error! Reference source not found.** for more information) and supporting data may be required for acceptance. Related to the EU REACH regulation, ECHA published the Read-Across Assessment Framework (RAAF) (ECHA, 2017b) in 2015, extended in 2017, to provide a framework and guidance for consistent evaluation of the scientific aspects of a proposed read-across case.

Although the applications of read-across may vary depending on the scope and different decision contexts, there is overall a common approach: The principle of the read-across technique is that endpoint information for one chemical (i.e. the source substance) is used to predict the same endpoint for another chemical (i.e. the target substance) lacking that information which is considered to be similar by scientific justification. The technique of read-across can be applied to characterise physicochemical properties, environmental fate, human health effects and ecotoxicity. For any of these areas, read-across may be performed in a qualitative or quantitative manner.

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

- One-to-one (one analogue used to make an estimation for a single chemical).

[42] Formalised system, usually computerised that enables an end-user to make rational predictions of toxicity based on structure alone. Expert systems are typically categorised by whether they are underpinned by: empirically based algorithms such as QSARs e.g. TEST, OPERA; knowledge bases such as SARs e.g. Derek Nexus, Toxtree or a hybrid of the two e.g. TIMES, ChemTunes.

[43] https://hpvchemicals.oecd.org/ui/ChemGroup.aspx

[44] https://www.oecd.org/en/topics/sub-issues/assessment-of-chemicals/integrated-approaches-to-testing-and-assessment.html

- Many-to-one (two or more category members used to make an estimation for a single chemical as part of a category approach).
- One-to-many (one analogue used to make estimations for two or more chemicals).
- Many-to-many (two or more category members used to make estimations for two or more chemicals as part of a category approach).

Read-across can be qualitative or quantitative. In qualitative read-across, the presence (or absence) of a property/activity for the target chemical is inferred from the presence (or absence) of the same property/activity for one or more source chemicals. Qualitative read-across gives a "binary" or "yes/no" answer. In quantitative read-across, the known value(s) of a property for one or more source chemicals is used to estimate the unknown value of the same property for the target chemical. Quantitative read-across is used to obtain a quantitative value for an endpoint, such as a dose-response relationship (e.g. NO(A)EL, LO(A)EL). Qualitative and quantitative read-across techniques are discussed in more detail in Section 3.2.2.

Structural similarity provides the most typical means of identifying likely source substances though custom searches factoring in other similarity considerations such as bioactivity similarity, physicochemical similarity, chemical reactivity, or metabolic similarity can also be undertaken. Structural similarity and similar properties and/or activities between the source and target chemicals are most often used as a basis for justifying read-across. Analogue suitability may be evaluated by reference to one or more of the following similarity contexts:

- Common functional group(s) (e.g. aldehyde, epoxide, ester, metal ion). An example is the ethylene glycols category assessed in the Cooperative Chemicals Assessment Programme[45]. Specific functional groups or substructures may be associated with specific reactive mechanism or metabolic pathways as defined in *in silico* profilers.
- Common constituent or chemical classes, similar range of carbon numbers. This is frequently the case with complex substances often known as UVCBs, for which information from starting materials/production process can also be taken into consideration.
- Common precursor and/or breakdown product that results via physical or biological processes (i.e. dissociation, metabolic or degradation pathway similarity). This is used to examine related chemicals, such as acid/ester/salt (e.g. esters of thioglycolic acid, thioglycolic acid and its ammonium salt). Additional examples are certain azo dyes based on carcinogenic components such as benzidine or other carcinogenic aromatic amines, where the carcinogenic aromatic amine is formed by the metabolism or degradation of the dye.
- Exposed constituent of a complex or multi-layered substance. Substances such as nanomaterials often have surface layers (coatings, surface modification) bound more or less strongly to its core. The core of the substance may be exposed to further interactions if the weakly bound surface modifier detaches. Otherwise, the chemical structure composing the surface modification layer may provide opportunity for read-across.

3.2.1. Hypothesis and evidence-based approaches

The process of developing an analogue or category starts with the identification of analogues and an evaluation of their relevance with respect to one or more similarity contexts. This is the overarching hypothesis to help substantiate any subsequent proposed read-across. The evaluation of analogues will include consideration of structure, composition, physicochemical properties, chemical reactivity, and ADME/TK (including metabolism), mechanistic and bioactivity similarity as discussed further in Section

[45] https://hpvchemicals.oecd.org/ui/ChemGroup.aspx

Error! Reference source not found.. This document details the elements that should be considered in hypothesis generation, in providing evidence, and gives examples wherever possible.

To increase confidence in the read-across approach when applied to analogues or a category, evidence should be provided to underpin the hypothesis on which the read-across is based. This can be done by adding new elements to reinforce and develop the initial hypothesis, or by providing new scientific evidence that the category parameter is behaving as expected. Perhaps the most compelling evidence in support of a read-across hypothesis is information on a common MOA of the substances and a mechanistic rationale for their common biological behaviour though it is recognised that this is likely to be possible in limited cases only. Evidence on the bioactivity similarity as generated from HTS/HCS or omics data will be more feasible as discussed in Chapter 2.

The hypothesis needs to be fit for a particular purpose and provide some mechanistic basis or understanding of why it is fit for that purpose. In many cases, use of predictions from (Q)SARs and expert systems may contribute to the hypothesis generation. This has been often applied using the QSAR Toolbox, where profiling results related to mechanistic endpoints are used in the category formation and read-across hypothesis. Careful evaluation of the rationale for all endpoints under consideration is needed to define the applicability domain of the analogue or category to ensure read-across is appropriate for the endpoints of interest. For example, if the hypothesis is based on structure and MOA, then it may be valid for certain aspects of mammalian toxicology but not hold for environmental endpoints. If the hypothesis is based on metabolism, then the read-across may only be valid for systemic mammalian endpoints or routes of administration, and local effects such as skin or eye irritation could be excluded. Interspecies differences in metabolism may need to be taken into consideration. In other words, the category members and the read-across hypotheses are endpoint-specific, although in many cases the same category members may be relevant for multiple endpoints. Once a hypothesis has been developed, it then needs to be examined using the available data (evidence) to see if the hypothesis is verified for the intended endpoint. This process of building a hypothesis and then testing with the available evidence can be referred to as the read-across justification.

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

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

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

3.2.2. Choice of qualitative or quantitative read-across

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

In deciding whether to use quantitative or qualitative read-across, the nature of the endpoint should also be considered. It may be expressed on a numerical or categorical scale. In most cases, a specific value is required for risk assessment, such as a No Observed Adverse Effect Concentration (NOAEC) or a NOAEL, an environmental half-life, or a partition coefficient. Illustrative examples of qualitative and quantitative read-across performed past and more recent are given in Table 4.

Challenges with read-across

An issue that may arise when read-across is carried out in the context of a category is that the experimental results for different category members may have been generated from studies that used different test methods or conditions (e.g. exposure route, duration) or species for a given endpoint. For example, in the case of reproductive toxicity, only screening studies may be available for some category members, whereas two-generation studies may be available for other members. Since the estimated results from the category approach have to be useful for risk assessment and classification, the uncertainty associated with the underlying study results has to be ascertained. It is clear that the scope of the estimated results for a member of a category should not exceed the scope of the underlying data for the other members of the category. As another example, for genotoxicity, if only *in vitro* results are available for some members of the category (source chemicals), only conclusions on *in vitro* genotoxicity might be reliably reached for the category members that lack data (target chemicals). If the scope of the underlying experimental results for an endpoint varies (e.g. a mix of results from screening tests and higher-tier tests), it is necessary to clarify the scope of the estimated results for the category members for which no experimental results are available. It may be possible to apply a WoE approach to all the data, which could lead to the same hazard identification for all the members of the category, irrespective of the data available for the individual compounds. However, there may be cases where there is sufficient information on the quantitative trend across categories that it may be desirable to segment the category and assign different hazard identification to subgroups within the category. In any case, the validity of the read-across and the hazard characterisation should be independent of the data requirements for the regulatory programme.

Qualitative read-across

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

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

If a regulatory classification is used to express the property or effect, differences in the potency of the chemicals could be sufficient to warrant different classifications, depending on the classification threshold. If a difference in the potency between source and target chemicals is suspected, for example based on trends in the available data, a quantitative read-across approach rather than a qualitative approach would

usually be required. This is particularly important where the target chemical is suspected to have a more stringent classification than the source chemical. A different classification can be considered where the classification criteria are based on the strength of the available evidence rather than a quantitative cut-off. In addition, differences in whether source and target chemicals cause direct or indirect effects may lead to differences in classification.

Quantitative read-across

In quantitative read-across, the known value of a property for the source chemical(s) is also used to estimate a quantitative value for the same property for the target chemical, which lacks data for the property of interest.

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

- Using the endpoint value of a source chemical, e.g. the closest analogue in a (sub)category[46].
- Using a trend to scale the available experimental results from two or more source chemicals to the target chemical[47] (it should be noted that the trend should be statistically sound), see Section 3.3.
- Processing the endpoint values from a few source chemicals e.g. by averaging from the nearest neighbours as it is applied in QSAR Toolbox or by averaging from N nearest neighbours by taking the most representative value, by computing a similarity weighted average as in GenRA.
- Taking the most conservative value of the closest analogues or the most conservative value in the (sub)category[48]. There is a possibility that the target could be more active than the closest analogue.
- Normalising target read-across based on molecular weight (MW) considerations[49] e.g. a molar adjustment of the lower 95% confidence limit of the effective dose (LED_{05}) as performed to account for differences in the molecular weights of 2-butanol and methylethyl ketone (MEK): 657 mg 2-butanol/kg-day x (72.1066 g MEK/mol ÷ 74.1224 g 2-butanol/mol) = 639 mg MEK/kg-day.

Which of the options is adequate to apply in a read-across depends on the read-across scope and decision context.

Quantitative read-across can also be used for complex substances/UVCBs, typically by applying data from substances with similar physicochemical properties (e.g. substances with similar boiling ranges, carbon ranges, composition) or by applying data from key/major constituents. However, quantitative read-across for a UVCB must be done carefully and requires a sufficient knowledge about the composition of the UVCB (i.e. sufficient knowledge about the identity, concentrations, and properties of its constituents) and hence an understanding of the key structures that are likely to drive the behaviour and properties of UVCBs. Hence read-across for UVCBs, if performed, may be more appropriate to be done in a qualitative or semi-

[46] For example, the OECD HPV Cooperative Chemicals Assessment Programme Gluconates category, where aquatic toxicity data for sodium D-gluconate were read-across to the calcium and potassium salts, D-gluconic acid and glucono-delta-lactone (Caley et al., 2007).

[47] For example, (OECD2006b) C6-22 Aliphatic Alcohols category, where internal QSARs were developed to predict aquatic toxicity based on K_{ow} and thus derive aquatic toxicities for the target chemicals.

[48] For example, the assessment within the EU Existing Substances Regulation and the OECD HPV Cooperative Chemicals Assessment Programme of zinc distearate used aquatic toxicity data from the more soluble zinc salts (chloride, sulphate) to derive the aquatic predicted no effect concentration ($PNEC_{aquatic}$) for zinc distearate (Tsakovska and Worth, 2007).

[49] https://iris.epa.gov/static/pdfs/0071_summary.pdf.

quantitative way than attempting to fill data gaps for employing quantitative approaches (quantitative read across, trend analysis, or QSAR approaches). This is further discussed in Chapter 6, Section 6.6.

Quantitative read-across can be used to fill data gaps in hazard and risk assessment. Assessment factors are often applied to toxicity values (e.g. NOECs from aquatic toxicity studies or LOAELs from repeated dose toxicity studies in rodents) to yield a dose or concentration to which humans or organisms may be exposed that is expected to be 'safe' (i.e. that does not result in adverse effects). However, the development or application of assessment factors when making read-across predictions is beyond the scope of this OECD guidance document. Application of assessment factors to reduce read-across uncertainty is highlighted in Chapter 2 under uncertainty considerations.

Table 4. Illustrative examples of quantitative and qualitative read-across

Target Chemical (CAS)	Source Chemical(s) (CAS)	Type of Read-Across	Purpose	Reference
Phenol, (1,1-dimethylethyl)-4-methoxy (25013-15-6)	4-tert-Butylphenol (98-54-4)	Qualitative	The source chemical was used to fill a data-gap for biodegradation. It was used as part of the WoE to support the half-life in water to evaluate environmental persistence in the ecological risk assessment	(Canada, 2010a)
Decanedioic acid, bis(1,2,2,6,6-pentamethyl-4-piperidinyl)-ester (41556-26-7)	Decanedioic acid, 1,10-bis(2,2,6,6-tetramethyl-4-piperidinyl) ester (52829-07-9) Decanedioic acid, 1-methyl 10-(1,2,2,6,6-pentamethyl-4-piperidinyl) ester (82919-37-7) Decanedioic acid, 1,10-bis[2,2,6,6-tetramethyl-1-(octyloxy)-4-piperidinyl] ester (122586-52-1)	Quantitative	Most conservative short-term oral value from a selection of analogues was used to calculate a margin of exposure for human health risk assessment	(Canada, 2010b)
Total Aluminum	Various aluminum-containing compounds	Semi-quantitative	43 studies on aluminum containing compounds used to characterize level at which neurological and reproductive/developmental effects begin to be repeatedly observed in animal studies	(Canada, 2010c)
MAPBAP acetate (72102-55-7)	5 (Q)SAR models + gentian violet (CAS 548-62-9), malachite green (CAS 569-64-2), C.I. Basic Violet 4 (CAS 2390-59-2) and leucomalachite green (CAS 129-73-7)	Qualitative	Identify potential health effects of target chemical to inform human health risk characterization	(Canada, 2010d)
Thiobis Propanoic Acid Derivatives	Propionic acid, 3,3-thiodi-, didodecyl ester (3,3-thiodipropionic acid, didodecyl ester) (CAS 123-28-4) Propionic acid, 3,3-thiodi-, dioctadecyl ester (3,3-thiodipropionic acid, dioctadecyl ester) (CAS 693-36-7) Propionic acid, 3,3-thiodi-, ditridecyl ester (3,3-thiodipropionic acid, ditridecyl ester) (CAS 10595-72-9)	Qualitative	Developmental toxicity testing in DLTDP (di-lauryl-thio-di-propionate) produced no adverse results in four separate species. In addition, in a 90-day repeat dose study with DLTDP, no effects on reproductive organs were observed. Since DLTDP is the smallest of the three materials it is estimated to be an appropriate conservative representative for the family	(EPA, 2001)

Target Chemical (CAS)	Source Chemical(s) (CAS)	Type of Read-Across	Purpose	Reference
Trimellitate Category	Tris-2(ethylhexyl) trimellitate ToTM) (3319-31-1)	Qualitative	Due to their higher molecular weight and bulky side chains, the remaining members of this category are expected to demonstrate a lower order of toxicity than ToTM. This is supported by a similar structural-activity relationship observed with phthalate ester compounds, i.e. the higher molecular weight phthalates (ester side chains >C7) are less active that the transitional phthalates (ester side chains C4-C6). Thus, the use of TOTM to represent the potential hazards of the other category members is a conservative position. No additional toxicity tests were proposed for this category	(EPA, 2007a)
Short chain chlorinated paraffins	Alkanes, C10-13, chloro-	Qualitative	Data gap filling. The NOAEL for effects via lactation was read across from medium chain chlorinated paraffins (both within the EU Existing Substances Regulation and the OECD Cooperative Chemicals Assessment Programme)	(OECD, 2000b)
P-t-butylphenol CAS 98-54-4	p-t-pentylphenol CAS 80-46-6	Qualitative	To flag a concern for further testing. Data from p-t-pentylphenol were used to request further testing on endocrine disruption in fish.	(EEC, 1993) and (Tsakovska and Worth, 2007)
Chlorobenzene-related chemicals	1,2,4-trichlorobenzene CAS 120-82-1, 1,2,4,5-tetrachlorobenzene CAS 95-94-3 (see Table 4, Section 4.2 of case study)	Quantitative	Prediction of NOAEL for 90-day rat oral repeated-dose toxicity for chlorobenzene-related chemicals	(OECD, 2020h)
Branched carboxylic acids	2-ethylbutyric acid CAS 88-09-51	Quantitative	Used a category approach to predict the outcome of a subchronic toxicity study for nine branched carboxylic acids	(OECD, 2020i)
Methyl hexanoic acid CAS 4536-23-6	See Table 1, Section 3 of case study.	Qualitative	Read-across based filling of developmental toxicity gap	(OECD, 2020j)
Saflufenacil CAS 372137-35-4	N-phenylimide pesticides	Qualitative	Read-across as part of a weight of evidence approach to fill data gaps/waiving rodent chronic toxicity and carcinogenicity bioassays	(Hilton et al., 2022) ; (OECD, 2024b)
Spiropidion CAS 1229023-00-0	Tetramic and tetronic acid pesticides.	Qualitative	Read-across as part of a weight of evidence approach to fill data gaps/waiving rodent chronic toxicity and carcinogenicity bioassays	(Hilton et al., 2022) ; (OECD, 2024b)
Bisphenols and their metabolites		Quantitative	A digitalized framework was established to enable quantitative read-across of a large number of substances and also allowing for human decisions for filtering and prioritization	(Yang et al., 2023)

3.2.3. Elements for a read-across justification

In developing an approach for data gap filling using either the analogue or chemical category, a number of general elements should be considered and discussed to demonstrate the relevance of the analogues such that the subsequent endpoint read-across is scientifically justifiable. No read-across justification is ever entirely identical because of the nature of substances and chemical classes, and the fact that these may be grouped by a variety of means, from similarity in structure, physicochemical properties, chemical reactivity profile, bioactivity, conventional toxicological profile and ADME/TK including metabolism, common toxicological or environmental properties, to common production, uses and applications (see Chapter 2.1.3).

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

- Chemical identity and composition:
  - Chemical structure e.g. SMILES, 2D structure, 3D structure.
  - Composition.
  - Impurities.
- Functional groups.
- Common substructures.
- Physicochemical properties and other molecular descriptors[50].
- Chemical reactivity e.g. use of structural alerts and *in chemico* assays.
- Kinetics: ADME.
- Bioactivity similarity as derived from data generated in HTS/TCS assays (e.g. ToxCast, omics studies).
- MOA/AOP.
- Chemical/biological interaction.
- Responses found in *in chemico* and *in vitro* assays.
- Information obtained from other endpoints/species/routes.
- The route and duration of expected exposure.
- Information on fate in the environment (environmental fate).

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

It should be noted that the order of this list does not necessarily imply a hierarchical approach, even though some hierarchy of considerations could be applied. For example, although structural similarity is very often a starting point for read-across, it may not always provide the best scientifically supportable basis for determining their relevance for group membership. However, the similarity in chemical structure is typically the first requirement to consider, assuming that the composition is analysed, and impurities or other

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

constituents are not expected to change the toxicity profile. Detailed explanation for how structural similarity determines the toxicity or the lack of it is particularly necessary for complex toxicity endpoints (e.g. cancer, reproductive toxicity) that are linked to multiple mechanisms. To the extent possible, the read-across should be done by linking structural moiety(ies)[51] in the source and target chemicals and their tautomers/isomers, metabolites, degradation products and derivatives to specific common mechanisms.

Chemical identity and composition

Chemical structure

The chemical structure of substances usually provides the initial rationale and impetus for developing an analogue or category approach. However, similarity of chemical structure hardly ever forms a complete justification for a category. Consequently, some categories that are constructed for purposes of assessment are likely to be composed of a subset of all the potential structures that could be envisaged. For example, OECD IATA Case Study No. 290 (OECD, 2018b) focused on a set of substituted phenols (hindered and non-hindered) and used an IATA to examine their estrogenic potential.

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

- Overall structural trend and/or structural similarity
- Functional group(s)/moieties
- Carbon chain length
- Linearity, branching
- Degree and position of branching
- Cyclics and aromatics
- Isomerism
- Stereochemistry e.g. 3D chemical structural information might be important for certain toxicological properties
- Salts and their relationship to a source chemical/parent substance
- Purity(ies)/impurities
- Metal speciation/valence

Composition

For some categories containing substances with multiple components/constituents[52] rather than single component/constituent, it is not possible to define the structures as detailed in the previous section. Multiple components/constituents may be introduced in the production process. On a case-by-case basis, and when available and not proprietary information, production process chemistry may inform the understanding of common structural elements within a category; similarly, variation in manufacturing processes may result in differences in chemical composition. Process chemistry may provide useful information on composition, purity and physicochemical boundaries, especially for process chemicals with

[51] In physical organic chemistry moiety is generally used to signify part of a molecule, (e.g. in an ester R1COOR2 the alcohol moiety is R2O). The term should not be used for a small fragment of a molecule (IUPAC, 2006).

[52] In the context of this guidance both terms can be used interchangeably.

a less-defined chemical structure. Requirements for reporting appropriate structural representations for polymers with no exact structural formulas should be specified, e.g. monomers, oligomers.

If the exact composition is variable or not known, these substances are typically referred to UVCB substances[53]. In the EU, substances in which more than one main constituent is present at a concentration ≥ 10% (w/w) and < 80% (w/w), with known composition, are referred to as multi-constituent substances (ECHA, 2023b).

Even though UVCBs vary in composition, it may be possible to define how UVCBs of a similar nature will act by demonstrating the similarity of composition between the source and the target chemicals. The challenge is that unlike categories of substances with defined chemical structures, UVCB categories are often composed of substances which do not have a single, well-defined structural difference between each category member. Rather, each UVCB substance can be visualised as more of a “cloud” or “sphere” of constituents with overlapping chemistries and properties, and there can be quantitative and qualitative differences in composition within the category. Consequently, a more expanded definition of the boundaries of their chemistries compared with the boundary definition of categories of discrete chemicals is a useful way to define a UVCB category. See for more guidance in Section 6.6.

It is not the purpose of this section to detail all the varieties of UVCBs and multi-constituent substances and how these might be addressed in detail. However, in building a category justification for read-across (ECHA, 2017c; ECHA, 2022b), it is necessary to be able to define the boundaries of the chemistry within the UVCBs so that some analysis can be made of the validity of the proposed read-across.

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

1. One approach is to conduct an assessment of the composition and trends of existing data among the category members, in order to understand the boundaries of the category. Within the intended UVCB or multi-constituent category, various components/constituents of the complex substances can be further assessed for any trends and the limits of those blocks defined. Such an analysis will yield whether target and source chemicals are indeed related and which ones will provide a valid read-across between one another. Which chemical blocks are used to help identify the category will be highly dependent on the nature of the substances involved. For some it will include specific chemical properties, and for others it will rely on blocks of similar structures.
2. A second approach is based on well-known hazards of constituents. For example, if hazards are known for components/constituents of the UVCB, then it may be possible to define which ones are of high hazard, and therefore would drive a hazard assessment. For example, if a petrochemical stream contains benzene or polycyclic aromatic hydrocarbons, recognised human carcinogens, their content in the UVCB would likely drive any human hazard assessment and there may be

[53] The range of different types of UVCB is very wide and the specific properties may be diverse, such that the applicability of a common approach needs justification. There are many different types of complex substances, although generally these all have the following characteristics in common:

- These contain numerous chemicals (typically closely related isomers and/or chemical classes with defined carbon number or distillation ranges) and cannot be represented by a simple chemical structure or defined by a specific molecular formula.
- These are not intentional mixtures of chemicals.
- Many are of natural origin (e.g. crude oil, coal, plant extracts) and cannot be separated into their constituent chemical species.
- The concept of “impurities” typically does not apply to complex substances in some OECD Member States.
- These are produced according to a performance specification related to their physicochemical properties.

limited value in considering any other information. Furthermore, if there are reliable data on the petrochemical stream as a whole, that data should be considered in a WoE.

Examples of categories and structural relationships

A number of UVCB categories have been assessed within the OECD Cooperative Chemicals Assessment Programme, three of which are described in Table 5.

Table 5. Examples of categories based on constituent/component analysis

Category	Constituents	Supporting information	Reference
C10-C13 Aromatics hydrocarbon solvents category	• Defined aromatic (single and double) hydrocarbon members with specific constituents / component profiles or composition • Defined boiling point range (manufacturing process defines product) • Carbon number range of category members identifies at minimum approximately 80% of the chemical constituents in the substance	• Carbon number ranges for all members • Boiling point ranges for all members • Maximum 10% naphthalene • Other purity criteria	SIDS Initial Assessment Profile (OECD, 2012b)
C14+ Aliphatics hydrocarbon solvents (<2% aromatics)	• Defined for members with specific constituents/component profiles or composition • Contains multi-constituent substances (UVCBs) that have a variable composition due to their chemistries and method of manufacturing	• Compositional analysis by: - Carbon number - Boiling point range - Maximum levels of aromatics, sulphur, and nitro compounds	SIDS Initial Assessment Profile (OECD, 2011c)
Anthracene oils	• Applicability – environmental endpoints only • Major constituents/components – various anthracene ring structures • Minor constituents/components – 3 to 4-fused aromatic sulphur-, nitrogen- or oxygen-heterocycles	• Main constituents/components and concentration ranges for all category members	Initial Targeted Assessment Profile (Environment) (OECD, 2009c)

Note: The selection of the examples in Table 5 was on an empirical basis and is intended to be illustrative of the types of chemistries and structures that have used the category approach for read-across. In the development of the highlighted categories other elements were also used in the justification for any read-across. The full category justification given in the reference should be consulted.

Characterising composition

When developing a category on the initial premise of composition for multi-constituent or UVCB category members, the compositional elements need to be stated in a clear and unambiguous manner together with the relationship between the various category members. The clearest way to do this is to build a table of the category members that indicates the elements that change and those that stay constant within the category and provide some indication of structure (see Table 6).

Table 6. Example for characterising category composition of UVCBs

Category member	Compositional Element 1 % Typical Concentration Range	Compositional Element 2 % Typical Concentration Range	Compositional Element 3 % Typical Concentration Range	Other common features from identification, production, or materials
1	10-20	60-80	10-20	Example: Produced by the reaction of R-x with y under condition of…
2	15-30	70-85	-	
3	60-80	10-20	0-10	

Common features may include such aspects as:

- Concentration ranges of a particular composition (e.g. typical ranges)
- Production orientated information:
  - Starting material
  - Boiling point ranges
  - Acid numbers
  - Carbon chain length

In some cases, a description of the production process can help establish the category and delineate its boundaries based on composition. Similarly, to reporting structural elements, the objective is to build an overall picture of the domain of the proposed category and identify or quantify to the best individual members, defining the relationships between its members and setting the boundaries of its chemical properties. In order to make the case for the category definition on compositional grounds, relevant supporting analytical data and physicochemical properties may be required to demonstrate how compositional properties change and bound the category.

Impurities

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

Consequently, when building a category, one must consider the issue of impurities and decide whether it is necessary to set limits on purity levels of the chemical(s) to ensure validity for any future read-across. A number of regulatory programmes of OECD Member Countries consider information on levels of impurities and how these affect intrinsic properties. The approach in setting purity limits is not consistent globally. As a related example, the EU Classification, Labelling and Packaging Regulation (CLP-Regulation[54]) includes provisions for the differential classification of substances based upon the presence of constituents with known hazards with specific concentration limits.

[54] Regulation (EC) No 1272/2008 of the European Parliament and of the Council of 16 December 2008 on classification, labelling and packaging of substances and mixtures, amended by Regulation (EU) 2024/2865 of the European Parliament and of the Council of 23 October 2024.

Functional groups

Common functional groups (e.g. aldehyde, ester, epoxide, metal ion) within a category are likely to be one of the critical elements that defines a category, and chemical structure also informs the likely physicochemical nature and chemical reactivity of the category members. Some functional groups are "structural alerts" which indicate the potential for a substance to cause toxicity through a particular mechanism that should be explicitly identified. Such groups may require more detailed examination and analysis through the use of *in silico* modelling tools.

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

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

Examples of categories and structural relationships

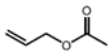
Many of the categories examined by the OECD and by Member Country regulatory programmes were initially conceived from structural considerations.

Selected examples of structural based categories of the Cooperative Chemicals Assessment Programme[55] and the OECD IATA Case Studies Project[56] are listed in Table 7. The list includes links to the published documentation that give the associated detailed rationale. It should be stressed that in all of these cases, the structural consideration was a starting point and not the entire justification. Further evidence to demonstrate why the read-across is valid between category members is always necessary. It should also be noted that this table is not an exhaustive list of categories conceived from structural considerations.

[55] https://hpvchemicals.oecd.org/ui/Default.aspx.

[56] https://www.oecd.org/en/topics/sub-issues/assessment-of-chemicals/integrated-approaches-to-testing-and-assessment.html

Table 7. Examples of structures and functional groups in selected categories from the OECD HPV Programme and IATA Case Study Project

Category	Structural relationship between category members	Functional groups	Number of substances in the category	Major justification for category	Reference
Alkyl chlorosilanes	Chlorosilanes, mono, di and tri	R-Si-Cl	4	Chemicals grouped based on similar molecular structure, high reactivity, physicochemical and toxicological properties	SIDS Initial Assessment Profile (OECD, 2010b)
Allyl ester	Esters of single allyl alcohol and saturated aliphatic carboxylic acid, C2-C18		19	Grouped based on the structure and metabolic hydrolysis of allyl acetate (C2) and allyl stearate (C18)	OECD IATA Case Study Project (OECD, 2016f)
Alkyl sulphates, alkane sulphonates and alpha olefin sulphonates	Alkyl sulphates with a predominantly linear alkyl chain length of C8-C18, C8-C18 alkane sulphonates, and alpha-olefin sulphonates with linear aliphatic chains of typically C14-C18	$R\text{-}SO_3^-$ cation	57 (43 Alkyl sulphates, 6 alkane sulphonates, 8 alpha olefin sulphonates)	Presence of predominantly linear aliphatic hydrocarbon chain with a polar sulphate or sulphonate group, neutralised with a counter-ion (i.e. Na+, K+, NH_4+, or an alkanolamine cation) is most important common structural feature. Close structural similarities result in physicochemical properties and environmental fate characteristic that follow regular patterns. Common physical and/or biological pathways result in structurally similar breakdown products and, with the surfactant properties, are responsible for similar environmental and very similar hazard profiles. Structural similarities result in same mode of ecotoxicological action. Varying length of the alkyl chain is most important parameter influencing ecotoxicity within each subcategory	SIDS Initial Assessment Profile (OECD, 2007a)
Alkylamidopropyl betaines	Amphoteric surfactants	Quaternary ammonium ion, carboxylic structure, amide bond		Category members are amphoteric surfactants containing a quarternary ammonium ion, a carboxylic structure, and an amide bond and are all manufactured from oils, usually coconut oil containing mixtures of C8 to C18 fatty acids and marketed as aqueous solutions (20-40%). Structural and functional similarities and comparable physicochemical properties of cocamidopropyl betaine inner salts and sodium salts suggest a similar ecotoxicological and toxicological profile. Values for physicochemical endpoints for lauramidopropyl betaine are similar or within the range of values for cocamidopropyl betaines, supported by accepted (Q)SARs, therefore similar ecotoxicological properties are assumed. All available physicochemical and environmental fate data are similar for lauramidopropyl betaine and cocamidopropyl betaine. MOA for aquatic toxicity should be the same because only the alkyl chain length differs for the chemicals in the mixture	SIDS Initial Assessment Profile (OECD, 2006c)

Category	Structural relationship between category members	Functional groups	Number of substances in the category	Major justification for category	Reference
Alpha olefins	Even-numbered, unbranched aliphatic chain (C6 – C14) with no other functional groups	alpha- olefin C_xH_{2x}		Category members are olefins bearing a single medium-length (C6–C14), even-numbered, unbranched aliphatic chain with no other functional groups. There is an increasing or decreasing trend or pattern from the shortest category member (C6) to the longest category member (C14) for various physical properties and ecotoxicity but there appears to be no difference across category members for biodegradation and health endpoints. Melting point, vapour pressure, and water solubility decrease with increasing chain length while boiling point and octanol:water partition coefficients increase with increasing chain length. Measured and predicted acute aquatic toxicity data indicate that 1-hexene, 1-octene, and 1-decene exhibit acute effects to aquatic organisms at levels at or below their water solubility, but 1-dodecene and 1-tetradecene are not likely to be acutely toxic. 1-hexene may be less toxic than the rest of the category members and 1-octene, 1-decene, and 1-dodecene are expected to be similarly toxic. Modelling could not predict the chronic aquatic toxicity of 1-tetradecene. No apparent difference regarding biodegradability. Data indicate no differences among the five category members for acute toxicity, repeat dose toxicity, genotoxicity, and reproductive/developmental toxicity	SIDS Initial Assessment Report (OECD, 2001a)
Aluminum alkoxides	Category members comprised of inorganic component and linear alcohol component		2	Category members have low toxicity to human health. Aquatic toxicity varies depending on the carbon chain length. All alcohols biodegrade and are not persistent	(EPA, 2010a)
Amine oxides	Alkyl hydrophobic substituent of different chain lengths with polar “head”	Amine oxide $R_3N\rightarrow O$	15	Category members have similar structures and functions. Substances are surfactants with a polar “head” (the amine oxide) and a relatively inert, hydrophobic “tail” (the long alkyl substituent). Structural variations are threefold: 1) the nature of the second and third substituents on the amine are either methyl groups or hydroxyethyl groups; 2) the long alkyl chain ranges in length from 8 to 20 carbons; and 3) the long alkyl chain may contain one or two double bonds. Alkyl chain lengths range from 8 to 20 with 12 and 14 being predominant. Average chain lengths for the mixtures are 12.9 to 13.5, with the exception of one tallow-derived compound. Presence of methyl- versus hydroxyethyl-substituents affects the basicity of the nitrogen only marginally, and the hydroxyethyl group lends more bulk to the hydrophilic head-group of the surfactant. Length of the longest alkyl substituent does not alter the chemical reactivity of the molecule but does affect its physical properties.	SIDS Initial Assessment Report (OECD, 2006d)

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

Category	Structural relationship between category members	Functional groups	Number of substances in the category	Major justification for category	Reference
	either 3-phenyl-2-propenal or 3-phenylpropanal backbone			features among members of this chemical category are that these contain either a 3-phenyl-2-propenal or 3-phenylpropanal backbone	
Dicarboxylic acid	Category composed of linear alkanes with common functional group (carboxylic acid) at each end of alkane chain		3	Category members have a common functional group, i.e. carboxylic acid, at end of alkane chain. Materials change by increase in carbon number from addition of CH_2 in alkane chain between carboxylic groups. Terminal carboxylic acids and limited chain length yield similar structure relationships. Members also share similar physicochemical, environmental fate, ecotoxicological, and mammalian toxicological properties	(EPA, 2001)
Ethylene glycols	Two terminal hydroxyl groups; the only variation is in the number of oxyacetylene units	glycols $HO(CH_2\ CH_2\ O)n\ H$, where n = 1-5	5	All category members have two terminal hydroxyl groups and differ from each other in the number of oxyethylene units. Members are therefore closely related in structure and have physicochemical properties that differ as expected resulting from increasing molecular weight and consistent functionality of hydroxyl moiety on each end of molecule. Hazard and dose response profile expected to change consistently, as confirmed by data and modelling	SIDS Initial Assessment Profile (OECD, 2004b)
Fuel oils	8 ethylene industry streams consisting predominantly of same higher-boiling hydrocarbons (mostly cyclic olefins and aromatics)		8 streams	Members are complex substances, containing variable amounts of alkanes, cycloalkanes, aromatics, olefins, asphaltenes, and hetero-molecules containing sulphur, oxygen, nitrogen, and organo-metals. Typically defined by process history, physical properties, and product use specifications. Streams that have undergone similar processing have similar physical/chemical/biologic properties and environmental fate and transport characteristics. Refinery streams within the heavy fuels category can therefore be grouped into seven subcategories based on their process histories	(EPA, 2014)
Long chain alcohols (C6-22 primary aliphatic alcohols)	Same basic structure $CH_3(CH_2)_nCH_2OH$ with variations with alkyl or methyl branching and some unsaturation in some of the 30 category members	R-OH	30	Category members share the same structural features, similar metabolic pathways, common mode of ecotoxicological action, and common levels and mode of human health related effects	SIDS Initial Assessment Profile (OECD, 2006b)

Category	Structural relationship between category members	Functional groups	Number of substances in the category	Major justification for category	Reference
Methyl xanthine	Tri or dimethyl xanthine	xanthine	3	Members share comparable structural, physicochemical, and molecular properties, and high level of structural similarity (all are methyl xanthines). Tanimoto score >0.9 based on CE-ToxGPS software	OECD IATA Case Study Project (OECD, 2020f)
Nitrates	Nitrate salts	Nitrate NO_3^-	7	Members are all inorganic salts which are solid under ambient conditions (except UAN (urea ammonium nitrate), which is a solution). Considered part of same category based on similar environmental fate, ecotoxicological and toxicological properties	SIDS Initial Assessment Profile (OECD, 2007b)
Substituted p-phenylenediamines	Category members phenylenediamines with various substituent groups always in para position of aromatic ring. Substituent groups may be all alkyl, all aryl, or mixed alkyl/aryl.		7	Category members share similar structures, physicochemical, and toxicological (including ecotoxicological) properties, and are not readily biodegradable	(EPA, 2001)
Sulphosuccinates	Category members have succinic ester backbone in which carbon alpha to one of the carboxyl functions has sodiumsulfo group in place of hydrogen atom		3	Category members have similar structures, follow a pattern regarding physicochemical properties and ecotoxicological endpoints, and share similar toxicological properties	(EPA, 2001)
Xylenes	Dimethyl benzene isomers	Bz–Me	4	Ortho- meta- and para-xylene are chemical isomers, and the only difference is the position of the methyl group on the benzene ring. Mixed xylene is a mixture of the three isomers and in addition, typically contains 15-20% ethylbenzene. Category members share similar physicochemical properties with the exception of the higher melting point of p-xylene. Toxicity of three individual isomers and mixed xylene is also similar	SIDS Initial Assessment Profile (OECD, 2003)

Category	Structural relationship between category members	Functional groups	Number of substances in the category	Major justification for category	Reference
C1 -13 Primary amines	Alkyl amines – Increasing alkyl chain and branching	primary amino-group RNH_2,	11	Category members are structurally similar with trends physicochemical properties and ecotoxicity and similar toxicological properties	SIDS Initial Assessment Profile (OECD, 2011b)
High molecular weight phthalate esters	Phthalic acid esters with carbon backbone R =>7	Ph-C-CO-O-R	7	Category consists of esters with alkyl carbon backbone with 7 carbon (C) atoms or greater. Category members contain linear and/or branched diheptyl, dioctyl, dinonyl, didecyl, diundecyl, didodecyl, and/or ditridecyl phthalate esters. Members also generally similar with respect to select physicochemical properties or display an expected trend. Members also similar regarding biological activity, i.e. these demonstrate few biological effects	SIDS Initial Assessment Profile (OECD, 2004d)

*References – Category SIDS Initial Assessment Profile OECD – Year and Link within the OECD Existing Chemicals Database https://hpvchemicals.oecd.org/ui/ChemGroup.aspx

Note: The selection of the examples in Table 7 was on an empirical basis from different regulatory and international programmes and is intended to be illustrative of the types of chemistries and structures that have used the category approach for read-across. In the development of the highlighted categories other elements were also used in the justification for any read-across. The full category justification given in the reference should be consulted.

Physicochemical properties

Physicochemical parameters are one critical determinant to the environmental and health properties of a substance affecting bioavailability, environmental fate, and thus the (eco)toxicity of a chemical. Consequently, the similarity (or logical trend) among the physicochemical properties of category members is an important element in building a read-across approach. If the source and the target chemical share similar properties, it might be hypothesised that there is a similarity of uptake and distribution in tissues of living organisms. Nevertheless, chemicals with equal physicochemical properties may still have different interactions with enzymes that could result in different metabolism and thereby distribution and elimination. The most used properties in this regard are shown in Table 8 below.

Table 8. Physicochemical properties important for hazard assessment

Property	Related to
• *log* K_{ow} • *log* K_{oa}	• Adsorption, estimation of bioconcentration in gill respiring animals, aquatic toxicity, mammalian absorption (oral and dermal) • Estimation of potential for bioaccumulation of non-metabolisable substances in air breathing animals
• Water solubility	• Adsorption, bioavailability, distribution in environmental media, Henry's law constant, aquatic toxicity, hydrolysis
• Molecular weight (MW) • Molecular dimensions	• Bioavailability, absorption or bioaccumulation, steric hindrance • i.e. 3 D structural characteristics such as D_{max} and molecular length (distribution or probability)
• Vapour pressure • Henry´s Law constant	• Volatility with respect to test conditions, inhalation • Distribution coefficient between air and water, potential for exposure from water-based formulation and hence relevant for considering inhalation route of exposure.
• Acid dissociation constant (*pKa*)	• Degree of ionization, relationship to irritation and corrosion, hydrolysis of ionisable substances, potential for uptake (including bioconcentration and accumulation), and sorption to soil (e.g. clay)
• *log D* (calculated)	• Lipophilicity, solubility, absorption, membrane penetration, plasma protein binding, distribution
• Boiling point • Melting point	• Informs on likely physical form at room temperature and pressure, potential exposure routes
• Surface tension	• Absorption - surfactants interact with the stratum corneum increase absorption of chemicals applied to the skin
• Particle size	• Depth of penetration of inhalable/respirable particles in the respiratory tract
• Viscosity	• Fluid behaviour and interactions in various biological system, ability to adhere to skin

For the environmental compartment in particular, the type of supporting information that is appropriate to report will depend on the environmental endpoint intended to be read across. However, basic physicochemical properties that determine environmental distribution and fate (e.g. MW, water solubility, partition coefficients such as *log* K_{ow}) will generally be useful. Particle size and structure are also relevant (e.g. nanoparticles).

For example, in the case of aquatic toxicity, similar *log* K_{ow} and aqueous solubility values between the source and target chemicals could be used to support the read-across, because *log* K_{ow} is known to be a determinant of the toxicity in aquatic organisms when the effect is mediated by mechanisms of narcosis. If the chemical is known or expected to act by a non-narcotic MOA, additional properties would provide useful supporting information.

While basic physicochemical properties (such as those listed in Table 8) are important factors in determining the boundaries of a given category, there may be practical problems for certain classes of chemicals such as gases, surfactants, PFAS, and especially for UVCBs. In these situations, consideration can also be made of alternative properties where these can be demonstrated to be applicable and robust

for such chemical classes (e.g. the membrane lipid-water partition/ distribution coefficient (K_{MLW}/ D_{MLW}) as an alternative for *log* K_{ow} for surfactants (ECHA, 2023c).

(Q)SAR predictions may be helpful if such a prediction provides a sufficient level of confidence to understanding potential trends within the category. It should be noted that these trends sometimes may not always be linear.

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

The similarity in physicochemical properties is very closely related to the trend analysis described in the preceding Section 3.2.2 for such endpoints and can be reported in a similar manner. However, the source of the information needs to be clear, whether the value is measured or calculated. Common software systems which contain models for predicting physicochemical properties include the QSAR Toolbox[57], Toxicity Estimation Software Tool (T.E.S.T.)[58], EPISuite™[59] and OPERA. Predictions from a range of different models are additionally available through the US EPA CompTox Chemicals Dashboard. The physicochemical properties are usually reported in a table. A plot of the trend or a boxplot depicting the distribution of the parameters across analogues or category members can be helpful to diagnose potential outliers. The pairwise comparison of the profile of different physicochemical properties across category members or between source and target substance can also be performed to derive a distance metric such as Jaccard, Euclidean or Pearson. Such comparisons to evaluate physicochemical profile similarity within categories or between target-source analogues have been described in (Yang et al., 2021) as well as in (Patlewicz et al., 2024b).

Absorption, distribution, metabolism, and excretion (ADME)

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

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

- The applicability of the available data from the source to the target substance.
- The need for supporting information on ADME (experimental and predicted) on both the source and the target substance to ensure that hypothesis is valid.
- The influence of variables in the category that are also changing, for example MW and *log* K_{ow}, and how these may influence read-across.

[57] QSAR Toolbox: https://qsartoolbox.org/

[58] T.E.S.T.: https://www.epa.gov/comptox-tools/toxicity-estimation-software-tool-test

[59] EPA EPISuite US EPA. 2012. Estimation Programs Interface Suite™ for Microsoft® Windows, v 4.11. United States Environmental Protection Agency, Washington, DC, USA.: https://www.epa.gov/tsca-screening-tools/epi-suitetm-estimation-program-interface.

- The applicability of the ADME data to the appropriate route of exposure(s).
- The applicability of the ADME data for specific endpoints.
- The applicability of the species in which the ADME data is obtained.
- Exposure route, including route-to-route extrapolation.

The Integrated Chemical Environment (ICE)[60] has tools for carrying out Physiologically Based Kinetic (PBK) analysis and chemical characterisation using supporting information which can provide additional insight into the chemical grouping exercise.

In addition, the use of *in silico* tools to simulate metabolites can be helpful in both the identification of potential analogues as well as in evaluating the metabolic similarity amongst candidate analogues. Various software tools exist that can simulate metabolism – from commercial applications such as Meteor Nexus, TIMES, CATALOGIC, to freely available tools such as the simulators within the QSAR Toolbox or the open-source Biotransformer tool (Djoumbou-Feunang et al., 2019). The scope and limitations of such tools are described further in Chapter 6. Various studies have explored how to apply such tools to the identification and evaluation of source analogues as part of an analogue or category approach. Examples include work by (Gadaleta et al., 2020) who outlined a framework to facilitate an automated evaluation of metabolic similarity in conjunction with physicochemical and structural similarity relying on a freely available metabolism prediction tool called SyGMa (Systematic Generation of Metabolites) (Ridder and Wagener, 2008). In a series of publications, the Laboratory of Mathematical Chemistry proposed a framework for assessing and evaluating metabolism similarity taking into consideration the identification of common metabolic pathway(s), common metabolite(s), and similarity of reactive metabolites formed. In (Yordanova et al., 2019) a methodical approach was described using the QSAR Toolbox where source analogues could be returned that took into account the metabolic activity pattern of the target substance. In (Yordanova et.al., 2021) the metabolic consistency between source and target substances was further developed based on side-by-side comparisons of metabolic maps. These considerations were subsequently implemented in TIMES and CATALOGIC (Kuseva et.al., 2021). In (Boyce et al., 2022), commonality in simulated transformations was used to compare candidate analogues for read-across. (Lester et al., 2023) defined metabolism fingerprints based on the transformations that are known to occur for the compound or predicted to occur based on transformation data for the functional group in a similar compound. A comparison (Tanimoto) was performed for the fingerprints defined for each target/analogue pair of structures.

Qualitative and quantitative characterisation of metabolic similarity

Current practice for assessing metabolic similarity within an analogue or category approach is largely limited to a qualitative inspection of the observed pathways and a subjective expert judgement assessment of the extent to which the transformations observed and potentially their rates are similar. An opportunity exists to leverage some of the same techniques that are used routinely in evaluating structural similarity and apply them to objectively evaluate metabolic similarity. The following approaches are outlined to highlight how such an evaluation might potentially be performed to address different aspects of metabolic similarity whether it be the similarity of the metabolites (i.e. the metabolites of the test chemical, which can also be referred to as (bio)transformation products), the similarity in their transformation pathways, the sequence of transformations, or the similarity in the overall metabolic graph (which describes a xenobiotic metabolic pathway, including chemicals and their connections). Application of these approaches is predicated on the availability of empirical or *in silico* predicted metabolism data.

[60] Integrated Chemical Environment (ICE) available on the National Toxicology Program website: https://ice.ntp.niehs.nih.gov/Tools.

- Evaluate the structural similarity of the metabolites produced from one or more source substances. This would involve characterising the metabolites by their structures and computing some type of chemical descriptor or fingerprint representation. The chemical representation could take the form of a bit vector of 1s and 0s denoting the presence or absence of certain structural features. A Tanimoto similarity index could then be computed to quantify the pairwise similarity assessment of a metabolite from parent A to that from parent B. This would quantify how similar metabolites from one target were to those metabolites from a source analogue.
- Evaluate the similarity of the presence/absence of metabolites produced from one or more source substances. Here a bit vector representation for the set of metabolites produced from the target/source analogues under consideration would be first derived. The vector would be labelled 1 to reflect presence of a specific metabolite versus 0 for absence of that metabolite. This type of representation would permit a comparison of the level of overlap between metabolites and represent that as a similarity metric. A high number between 0 and 1 would reflect a high degree of common metabolites between target and source substances. Alternatively, the counts of metabolites could be computed to reflect the frequency of metabolites being produced from a target substance relative to its source analogue. This was an approach applied in (Boyce et al., 2022). A different approach could be to use SMARTS or substructures including metabolic SMARTS for the comparison of metabolism. A recent example of this has been used in a category approach for the strobilurene fungicides and sulfonyl urea herbicides (Enoch et al., 2022; 2023).
- Evaluate the similarity of transformations produced from 1 or more source substances. Create a bit vector representation of all unique transformations produced by a given target or source analogue. The vector would be labelled 1 to reflect a particular transformation occurring for a given parent substance whereas 0 would indicate that transformation was not expected to occur. A similarity metric such as the Tanimoto index could be computed to reflect the degree to which the set of transformations between one chemical and another was similar. This would provide a quantitative metric of the degree of commonality in transformations observed. An example of this approach was applied in (Lester et al., 2023).
- Evaluate the similarity in the metabolic graph produced from 1 or more source substances. There are a number of approaches of characterising the overall metabolic graph for a given chemical. Much research over many years has been devoted to the area of measuring the similarity of two graphs. Approaches include distance measures, graph kernels as well as graph embeddings. For more information, the reader is referred to the following review articles that provide some context for these approaches – see (Ma et al., 2021) and (Wills and Meyer, 2020). This type of evaluation would provide some insight as to the level of similarity in the structure of the metabolic graph and its sequence of transformations.
- Evaluate the similarity between metabolites produced by two or more compared chemicals based on three individual similarity criteria: (1) commonality of the metabolic transformations, (2) structural or reactivity pattern of the produced metabolites, and (3) common metabolites (Yordanova et al., 2021; Kuseva et al., 2021). This approach is automated in TIMES and CATALOGIC platforms[61], with each metabolic similarity criteria applied together having additional settings and a weight of contribution; a fingerprint and a hologram comparison can be used to evaluate the similarity.

For a metabolic category, it is hypothesised that a substance A is metabolised to a series of other substances and therefore the hazard data from the substance A can be used to identify the hazard of the metabolites and vice versa. This rationale is discussed in more detail in Chapter 6. It is likely that such a category is limited to a set of primary metabolites, because primary metabolites are likely to act more similarly to the parent compounds than are more distant (e.g. secondary) metabolites. In addition, Phase II metabolites (conjugates) are generally accepted as not toxic and readily excreted, although they can be

[61] https://oasis-lmc.org/products/software.aspx

further metabolised (bioactivated) into toxic metabolites, for example in the kidney. However, care should also be taken when reading across among the parent and primary metabolites. For example, it may not be useful to use data derived only from a single primary metabolite to represent toxicity of the parent compound (and/or other metabolites) because such data would not reflect the additional activity/toxicity associated with other metabolites of the parent chemical, or the parent chemicals themselves. Kinetics should consider exposure duration of the parent and or metabolite(s) which may be critical for the possibility of employing the read across approach. When knowledge about an endpoint of concern for one metabolite is available and not for the parent compound, it is helpful that the exposure time to the parent compound compared with that of the metabolite is very short in order to use the data available for the metabolite to read-across to the parent compound. It should be plausible that such a short exposure to the parent compound would not lead to effects for the endpoint in question. How short can depend on the endpoint under consideration, the transformation concerned e.g. a well-established hydrolysis reaction as well as the level of uncertainty that can be tolerated for the decision context in hand.

If data for a single metabolite are to be used, it is necessary to build the argument, not only on the metabolism data but on other criteria as well, such as knowledge that the metabolite is the only metabolite associated with toxicity or that other metabolites are likely to contribute only slightly to the overall toxicity. In such cases, additional information may also need to be considered, such as chain length and common functional groups. The chain length may result in lack of metabolic transformation (for this reason the octylchlorosilane was not included in the OECD Alkyl chlorosilanes category, (OECD 2010b)), or change the metabolism (for this reason the tributylamine was not included in the OECD Tertiary amines category, (OECD,2012c)). Addition of similar substances with analogous data would also broaden the general confidence in the approach. For example, two analogue pairs, each comprising an ester and its corresponding alcohol was used to strengthen the read-across of the 90-day endpoint being proposed in (Ball et al., 2014). Further information on metabolic categories is discussed in Chapter 6.

Table 9. Examples of categories with similarities in ADME/bioavailability

Category	ADME	Common structural features	References
Cyclohexyl Derivatives Category	4-tert-butylcyclohexyl acetate will undergo hydrolysis to yield 4-tert-butylcyclohexanol. Subsequently 4-tert-butylcyclohexanol is conjugated with glucuronic acid to yield the corresponding glucuronide that is excreted mainly in the urine.	This category consists of 2 substances, 4-tert-butylcyclohexanol and its corresponding acetate ester, 4-tert-butylcyclohexyl acetate	(EPA, 2007b)
Pyridine and Pyridine Derivatives	Both piperidine and pyridine are readily absorbed through the gastrointestinal tract, skin and lungs, and eliminated primarily via the urine. Although these do not have a common metabolite, both chemicals have been shown to undergo metabolism via C- oxidation and N-oxidation, and N-methylation has been shown to be a metabolic route for pyridine. Therefore, piperidine would be expected to be metabolized and eliminated in a similar manner and rate as pyridine.	All members of the Pyridine and Pyridine Derivatives Category are structurally-related derivatives of pyridine in that these are based on the pyridine unsaturated ring structure. Piperidine (CAS RN 110-89-4) is simply the saturated ring structure derivative of pyridine.	(EPA, 2004a)
Terpenoid Primary Alcohols and Related Esters	Geranyl acetate is rapidly hydrolysed. The alcohols geraniol, nerol, and citronellol are efficiently detoxicated by two principal pathways. In one route, the alcohols are successively oxidized to the corresponding aldehydes and carboxylic acids, the latter of which are selectively hydrated or reduced. In a second route, the aldehydes undergo reduction to the corresponding alcohols that are substrates 11 for omega-oxidation to eventually yield diacids and their reduced or hydrated analogues. Polar metabolites formed via these two pathways will be efficiently excreted primarily in the urine as the glucuronic acid conjugates.	Citronellol, geraniol, and nerol are close structural relatives. Nerol and geraniol are cis/trans isomers of 3,7-dimethyl-2,6-octadien-1-ol and citronellol is the dihydro analogue of geraniol (3,7-dimethyl-6-octen-1-ol).	(EPA, 2004b)
Sulfosuccinates	It is likely that these esters will be metabolized in rodents by esterases. Compounds formed from deesterification will be similar for all three molecules, with the exception of the alcohol moiety. Whereas de-esterification of sodium diethylhexyl sulfosuccinate gives rise to 2-ethylhexanol, similar metabolism of sodium dicyclohexyl sulfosuccinate leads to the formation of cyclohexanol. Likewise, metabolism of sodium 1,3-dimethylbutyl sulfosuccinate leads to methyl isobutyl carbinol.	The general structure for the category is defined as dialkyl sodium sulfosuccinate or dicycloalkyl sodium sulfosuccinate. This describes a molecule witha succinic ester backbone, in which a carbon alpha to one of the carboxyl functions has a sodiumsulfo group in place of a hydrogen atom.	(EPA, 2002a)
Phosphoric Acid Derivatives	Metabolism studies conducted on the tributyl phosphate indicate that dealkylation to form the alkyl alcohol is the primary route of metabolism. The phosphoric acid tri-esters are rapidly metabolized to di-esters with mono-diesters also being produced. Studies of tributyl phosphate show that 40-64% of the parent compound is metabolized to dibutyl dihydrogen phosphate and that 1 1 -2 1 % is metabolized to the monobutyl species. Therefore, tris(2-ethylhexyl) phosphate is expected to be metabolized to bis(2-ethylhexyl) phosphate (CAS# 298-07-7) and mono(2-ethylhexyl) phosphate (CAS# 1070-03-7). Based on the evidence for dealkylation as the primary metabolic pathway, 2-ethylhexanol is the expected metabolite of tris(2-ethylhexyl) phosphate (CAS# 78-42-2) and 2-ethylhexyl phosphate (CAS# 12645-31-7). Triisobutyl phosphate is expected to be metabolized similarly as tributyl phosphate, with methoxypropanol as the alcohol metabolite.	The chemicals within the category are defined as esters of phosphoric acid, having a phosphoric acid backbone with various alkyl substituents as illustrated.	(EPA, 2004c)
Cinnamyl Derivatives	The aromatic cinnamaldehyde derivatives are readily oxidized to cinnamic acid derivatives. The urinary metabolites of cinnamyl alcohol and cinnamaldehyde are mainly derived from metabolism of cinnamic acid.	The four substances in this group are un-substituted or alkyl-substituted cinnamaldehyde or 2,3-dihydrocinnamaldehyde derivatives. Common structural features among members of this chemical category are that these contain either a 3-phenyl-2-propenal or 3phenylpropanal backbone.	(EPA, 2005)

Category	ADME	Common structural features	References
Benzyl Derivatives	The benzaldehyde derivatives are readily oxidized to the corresponding benzoic acid derivatives while the benzyl esters are hydrolyzed to yield benzyl alcohol that is subsequently oxidized to benzoic acid as a stable metabolite or endproduct. The benzoate and 2-hydroxybenzoates esters are hydrolyzed to yield benzoic acid and 2-hydroxybenzoic acid derivatives, respectively. The benzaldehyde derivatives are readily oxidized to the corresponding benzoic acid derivatives while the benzyl esters are hydrolyzed to yield benzyl alcohol that is subsequently oxidized to benzoic acid as a stable metabolite or endproduct. The benzoate and 2-hydroxybenzoates esters are hydrolyzed to yield benzoic acid and 2-hydroxybenzoic acid derivatives, respectively. As a stable animal metabolite, benzoic acid derivatives are efficiently excreted primarily in the urine.	The 10 substances are placed in the same category because all contain a benzene ring bonded directly to an oxygenated functional group (aldehyde or ester) that is hydrolyzed and/or oxidized to a benzoic acid derivative."	(EPA, 2002b).
TDAs	3,4-TDA (3,4-Toluenediamine) was assessed for the derivation of screening level Provisional Peer Reviewed Toxicity Values (PPRTVs). Other TDAs were identified with authoritative repeated dose toxicity data and an evaluation of analogue suitability was made using not only structural considerations, but also with respect to toxicological, metabolic and reactivity similarity. The use of TIMES and Meteor was used to assist in the assessment of metabolic similarity in conjunction with available *in vivo* data.	4 positional isomers to 3,4-TDA were identified. Major metabolic steps for the TDA analogues were acetylation of amino groups, ring hydroxylation with some evidence for oxidation of the methyl groups. Predicted metabolic pathway transformations were consistent with the reported pathways where available.	(EPA, 2021)
Allyl ester	In rats, allyl esters are hydrolyzed by hydrolytic enzymes to allyl alcohol and carboxylic acids in the intestine, liver, and/or other tissues (Silver and Murphy). Following hydrolysis, the liberated allyl alcohol is distributed in the liver, and then predominantly oxidized by hepatic alcohol dehydrogenase (ADH) to acrolein, which is further oxidized by aldehyde dehydrogenase (ALDH) to acrylic acid. Acrolein forms a conjugate with glutathione (GSH) in the liver. The conjugate is finally excreted in urine as 3-hydroxypropylmercapturic acid (3-HPM) through metabolic processing.	Grouped based on the structure and metabolic hydrolysis of allyl acetate (C2) and allyl stearate (C18).	(OECD, 2016f)
Dipropylene glycol methyl ether acetate (DPMA)	DPMA and its source analogue Dipropylene glycol methyl ether (DPM) were grouped together on the premise that DPMA was readily hydrolysed by esterases to its corresponding alcohol DPM. A commensurate analogue pair of Propylene glycol methyl ether (PM) and propylene glycol methyl ether acetate (PMA) provided supporting evidence for the rate of the hydrolysis reaction and the paired toxicity profile across endpoints including the repeated dose toxicity (90 study) and screening level prenatal developmental toxicity being read across.	The four substances in this group were 2 analogue pairs of methyl ether and their acetates. Their underlying grouping rationale was the hydrolysis of the acetate group to alcohol.	(Ball et al., 2014)
alpha-terpinyl acetate and alpha-terpinyl propionate	Alpha-terpinyl acetate and propionate are esters of the alpha-terpineol and have very similar features, reactivity, physicochemical properties, and metabolic pathways. The likely metabolic pathways for both substances involves conversion (hydrolysis) to alpha-terpineol, which would be the key toxicological species, and the simple acids (acetic and propionic). Consequently, the metabolism will not diverge in a manner which could lead to different toxicological outcomes.	alpha-terpinyl acetate and alpha-terpinyl propionate have one methyl group difference in their chemical structures	(Wu et al., 2010)
dodecanoic and stearic esters of sorbitol	Dodecanoic and stearic esters of sorbitol are potential target and analogue substances. These have similar structural features, reactivity, and metabolise in a similar manner - hydrolysis of the esters to release sorbitol and their corresponding fatty acid. However, compared to the dodecanoate sorbitol, the stearate sorbitol has a longer alkyl chain that affects its physicochemical properties. The estimated *log* K_{ow} value of the stearate ester is three units higher than for the dodecanoate ester. Respectively, the water solubility is lower for the longer-chain ester. The difference in the chain length should not have a significant effect on the metabolism pathway and the mode (and/or mechanism) of action. However, the bioavailability depends on solubility and	Dodecanoic and stearic esters of sorbitol share a common scaffold, but their chain length differs to the extent it can affect its absorption characteristics.	(Wu et al., 2010)

Category	ADME	Common structural features	References
	thus on the length of the alkyl chain. The two substances would be expected to have quite different absorption rates through the dermal route, but absorption during ingestion would not be expected to be so different. Thus, consideration would be needed, depending on the route of exposure and the toxicological endpoint of interest, to incorporate this added uncertainty.		
4-alkylsubstituted phenols	4-alkylsubstituted phenols, which can be oxidised to form quinone methide derivatives and may generate similar toxic effects such as cytotoxicity and skin sensitisation, caused by the (bio)transformation product.		(Wu et al., 2010)

Chemical reactivity

A key aspect in comparing a target and source substance is in terms of its chemical reactivity. This can provide insights into likely transformations that might occur or how a specific toxicity might be mediated. Wu et al. (2010) outlined several practical ways of evaluating reactivity considering the following features:

- Similarity in structural alerts. Structural alerts encoding electrophilic reactivity such as the profilers that exist within the QSAR Toolbox provide a convenient means of assessing whether target and source substances share common features such as nitrosamines, quinones etc. that have been associated with specific toxicity effects. SMILES arbitrary target specification (SMARTS) features encoded in Toxtree, or OCHEM ToxAlerts are just a couple of examples of other collections where substances can be potentially profiled on the basis of structural alerts. (Nelms et al., 2019) reviewed the commonality of some of these alert schemes. In addition to structural alerts as encoded *in silico* tools, there are also *in chemico* and HTS assays available to measure electrophilic reactivity. Notable examples are the glutathione assay developed by (Schultz et al., 2005) as well as the MSTI ((E)-2-(4-mercaptostyryl)-1,3,3-trimethyl-3H-indol-1-ium) assay by (McCallum et al., 2013). The MSTI assay was applied to the Tox21 library in (Patlewicz et al., 2023) and the results compared to the established structural alerts.
- Similarity in functional groups. Certain functional groups e.g. aldehydes, esters, ketones and core scaffolds e.g. phenyl rings, alkyl chains will play a role in determining chemical reactivity and potential sites for metabolic transformation.
- Similarity in position of double bonds. Presence and position of double bonds can impact both the reactivity and toxicity. Conjugated double bonds or proximal to functional groups can contribute to increasing the reactivity e.g. alpha, beta-unsaturated aldehydes. Some of these double bond considerations are already captured within published structural alerts.
- Presence of additional functional groups. Other functional groups besides those driving the reactivity may play a role in modulating the reactivity potential. For example, steric hindrance can reduce reactivity.

Chemical/biological interaction

The interaction of a chemical at a molecular target leading to a particular adverse outcome, also called the molecular initiating event (MIE), can occur via different mechanisms. Many MIEs are defined in the form of covalent binding to proteins and/or DNA. These types of MIEs are based on the principles of reactivity of organic chemistry (i.e. electrophile-nucleophile reactivity). In contrast, 'receptor binding' or binding to enzymes are often based on non-covalent interaction, which are more selective in nature. For these interactions, the 3D structure of the molecules can be important (see examples from OECD IATA Case Studies (OECD 2020k; 2020l) and (van der Stel et al., 2021)). Within the AOP concept the MIE represents a primary anchor as the beginning of the cascade which can be linked to the intermediate key events leading to the specified final adverse outcome.

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

For several endpoints, such as skin sensitisation or mutagenicity, knowledge about chemical reactivity provides important information when applying the analogue and category approach. For skin sensitisation, one of the necessary steps a chemical has to undergo is to form a stable association with a skin protein (OECD, 2012a). This is thought to be a covalent association where the chemical behaves as an

electrophile and the protein as a nucleophile. A similar analogy is relevant for gene mutagenicity but where DNA represents the nucleophile. Structural alerts as encoded in Toxtree, TIMES, Derek Nexus, and the QSAR Toolbox could be used to characterize the specific reaction mechanism leading to covalent binding to protein or DNA. Experimental "*in chemico*" systems (OECD, 2025a) could also be used to quantify the electrophile-nucleophile reactivity and to confirm the predicted reaction mechanism to support a read-across for e.g. skin sensitisation (Aptula et al., 2006) or mutagenicity (Benigni et al., 2005).

Biological read-across

There is increasing use of molecular screening approaches, such as *in vitro* assays coupled with omics technologies, HTS and HCS, (see Section 2.4.2) to inform biological activities/endpoints of chemicals. Many new assays are being used to screen large numbers of chemicals for particular effects or used to characterise the intermediate key events within an AOP.

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

In order to provide solid mechanistic reasoning to use *in vitro* methods, it is useful to have transparent descriptions of a plausible progression of effects at the different levels of biological organization provided by AOPs. As explained in Section 2.4.3, the AOP approach is a bottom-up approach where events measured at the *in chemico* and *in vitro* level are linked to events measured at the *in vivo* level. For example, in fish, estrogen agonists bind to the estrogen receptor, which can be measured *in chemico*[62] and *in vitro*, and set off a cascade of responses including the up regulation of vitellogenin production in the liver, which can also be measured *in vitro*, the conversion of testes to ova and the feminization of males observed *in vivo* leading to reproductive impairment and a decrease in the population.

The incorporation of *in vitro* methods into any integrated testing or assessment scheme allows for the employment of relatively rapid and often but related to the complexity of the AOP and number of *in vitro* tests needed to cover all significant MIE(s) and KEs, inexpensive hypothesis-driven testing. The same goes for employment of *in chemico* methods – and for predictive models building upon training set data generated by use of such *in vitro* and *in chemico* methods. In such a scenario, the hypothesis that the target and source chemical have similar adverse effect on an apical endpoint can be tested by applying appropriate *in silico*, *in chemico*, and *in vitro* methods identified from the integrated scheme and depending on their reliability.

Bioactivity similarity

In many cases, no AOP may exist for a specific endpoint to provide the mechanism/mode of action information to substantiate a read-across. Indeed, the available profilers for specific MIEs are mostly limited to endpoints such as skin sensitisation and mutagenicity where the biological mechanisms are well established. In these scenarios, high throughput/high content screening or omics data can provide compelling evidence of similarity in biological responses.

In Chapter 2, a roadmap of the different ways in which HTS/HCS or omics data could be utilised were outlined depending on the type of data, the extent to which it could be directly associated with a specific endpoint or was associated with mechanistic evidence. Analogues can be identified for a target chemical on the basis of such data. The assumption in this case is that the target chemical has been tested in an array of HTS/HCS/omics assays and the profile is represented as a 'fingerprint' of hitcalls or potency measures. This fingerprint might be constrained in representing a defined signature or be more general to

[62] *In chemico* refers to an abiotic measurement of chemical reactivity.

capture a broad profile of responses e.g. a high throughput phenotypic profiling fingerprint versus a targeted metabolomic signature. In either scenario this fingerprint can then be compared against an inventory of potential source analogues that have been tested in the same suite of assays and therefore are associated with the same fingerprint representation. A pairwise similarity metric would return the top n analogues on the basis of the similarity in biological profile. Such a construct has been developed as part of the Generalised Read-Across (GenRA) application (Patlewicz and Shah, 2023; https://comptox.epa.gov/genra/) wherein a target chemical with associated ToxCast data can be used to profile against other substances tested in the ToxCast suite of assays to return analogues similar with respect to their ToxCast hitcall outcomes. The same approach is not restricted to ToxCast data, indeed in (Tate et al., 2021), a systematic data-driven analysis was performed to evaluate the role that a targeted transcriptomic dataset could be used to search for similar analogues. Related efforts to investigate the use of transcriptomic data for read-across have been undertaken by (DeAbrew et al., 2019) and (Wu et al., 2023) and discussed more broadly by Viant et al. 2024a). The use of the biological fingerprint can also be used to compare the bioactivity similarity of candidate analogues identified on the basis of structural similarity. Use of biological fingerprints in this manner are particularly helpful where the linkages to apical endpoints or other KEs in an AOP are not always apparent.

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

In analysing the elements of read-across justification, mode/mechanistic understanding is a key element. The term “mode of action” is understood in a broader sense than “mechanism of action”, the first being seen as an integrator of the general type of interaction of a chemical with the organism, while the second is perceived as the precise (bio)chemical molecular interaction related to the Molecular Initiating (MIE) or Key Event (KE) of an AOP. An AOP, for any given hazard endpoint, can be the basis for developing an IATA (OECD, 2017g). With each grouping a description of the likely mode or mechanisms of action is specific and should be considered together with its limitation and purpose.

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

The mechanistic basis of developing a category including modes and/or mechanisms of action[63] is described in Chapter 2, specifically Section 2.4.3 that includes discussion of the development of AOPs.

If it can be demonstrated that the mode or mechanism of action for the toxicological or ecotoxicological effect is the same for similar structures or functional groups, then the confidence of the read-across from a source to a target chemical is significantly increased. However, it should be mentioned that having the same MOA does not necessarily mean that the toxicity will be reliably estimated as variation of toxicity within each MOA might exist.

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

[63] Defined in Chapter 2.

With the increasing availability of mode of action information, especially through omics technologies (e.g. transcriptomics, metabolomics) and other high content screening (HCS), high throughput screening (HTS), AOPs, as well as other predictive data, more integrative approaches such as IATA can be explored to develop and/or support hypotheses/justifications for read-across and selection of the most appropriate molecular descriptor(s)[64]. In one study (van Ravenzwaay et al., 2016), metabolomics information together with the evaluation of the classical toxicological parameters from a 28-day study formed the basis of a substantiated claim to waive the 90-day study for the selected compound if the reference compound(s) is/are convincingly similar. Lizarraga et al. (2019) applied an expert-driven read-across approach to identify and evaluate analogues to fill non-cancer oral toxicity data gaps for p,p'-dichlorodiphenyldichloroethane (p,p'-DDD). The source analogue p,p'-dichlorodiphenyltrichloroethane (DDT) was proposed. Among the primary similarity contexts (structure, toxicokinetics, and toxicodynamics), toxicokinetic considerations were instrumental in separating p,p'-DDT as the best source analogue from other potential candidates (p,p'-DDE and methoxychlor). However, *in vitro* HTS assays from ToxCast were also used to evaluate similarity in bioactivity profiles to build additional confidence in the read-across approach.

Mode and mechanism of action concepts can also facilitate the read-across for aquatic toxicity. According to one of the earliest classification schemes (Verhaar et al., 1992), four modes of actions are distinguished for acute aquatic toxicity: inert substances, relatively inert substances, reactive substances, and specifically acting substances. The toxicity of the substances in the first two groups (later known also as non-polar and polar narcotics, respectively) is mainly hydrophobicity driven, while the second two groups (i.e. the reactive chemical substances and the specifically acting substances) form specific domains, and read-across between such domains is not trivial. The more precise definition of the mechanisms of aquatic toxicity can further facilitate the filling of data gaps. Some authors distinguish, instead of between reactive and specifically acting substances in relation to fish, between uncoupling of the oxidative phosphorylation, respiratory inhibition, and electrophilic/nucleophilic mechanisms, electrophiles/proelectrophiles, acetylcholinesterase inhibitors, or central nervous system seizure agents (Russom et al., 1997). Other authors split further the electrophilic reactivity in specific reactivity mechanisms such as Michael type-addition, Schiff-base formation, etc. (Schultz et al., 2005). Different types of models could be used within a specific mechanistic domain (Netzeva et al., 2008). For substances within the same reactive mechanism of action, the potency of protein binding as predictor for e.g. acute aquatic toxicity, can be estimated in (semi-)quantitative manner (QSAR Toolbox).

[64] see the OECD IATA Case Studies Project: https://www.oecd.org/en/topics/sub-issues/assessment-of-chemicals/integrated-approaches-to-testing-and-assessment.html

Table 10. Examples of categories with similarities in MOA

Category	MOA	Common structural features	Reference
Mononitroanilines	Toxicity is characterized by the ability to form methemoglobin in both humans and animals.	The chemicals selected for inclusion in this category are isomeric forms of the same base chemical	EPA (2003a)
Fuel Oils	The aquatic toxicity of products in the category are expected to fall within a narrow range regardless of the varying carbon number range and constituent composition of those products, because the constituent chemicals of those products are neutral organic hydrocarbons whose toxic mode of action is non-polar narcosis. The mechanism of short-term toxicity [fish] for these chemicals is disruption of biological membrane function, and the differences between toxicities (i.e. LC/LL_{50}, EC/EL_{50}) can be explained by the differences between the target tissue-partitioning behaviour of the individual chemicals.	The category was developed by grouping 8 ethylene industry streams made up of hydrocarbons that are generally carbon number 8 (i.e. C8) and higher with varying amounts of lower boiling materials. The streams are similar in that these are all complex streams that consist predominantly of the same higher-boiling hydrocarbons	EPA (2003b)
Allyl ester	Following the hydrolysis of allyl esters, allyl alcohol is readily oxidized to acrolein by ADH in the liver. Acrolein is a highly reactive substance that appears to cause hepatotoxicity. It readily forms an acrolein-GSH adduct, leading to GSH depletion, oxygen radical formation, and lipid peroxidation. Acrolein is also capable of reacting with cellular macromolecules nonenzymatically via Michael additions. Reactions with critical intracellular proteins and subsequent adduct formation are proposed as one component of the cytotoxicity of acrolein. Additionally, it has been proposed that oxidative stress subsequent to the loss of GSH may be related to mitochondrial dysfunction. These biochemical events caused by acrolein are believed to be associated with hepatocellular damage and death.	Esters of single allyl alcohol and saturated aliphatic carboxylic acid, C2-C18.	(OECD, 2016f)
Phenoxy herbicides	The metabolome evaluation of the source substances indicates liver and kidney as the target organs. The metabolome profile associated with the liver indicates a lipid reducing activity comparable with the one induced by peroxisome proliferators. A hallmark of phenoxy herbicides is saturation of renal excretion and reabsorption causing a rise in compound blood levels associated with the onset of renal toxicity.	Chemicals selected for this study share phenoxypropionic acid substructure.	(Ravenzwaay et al., 2016)
PFAS	All the PFAS tested in this study induced transcriptional changes in cholesterol biosynthesis and lipid metabolism pathways, and predicted PPARα activation. The data indicate that these PFAS may have common molecular targets and toxicities.	Per-and poly-fluoroalkyl substances with 4-10 carbon chains.	(Rowan-Carroll et al., 2021)

3.3. Trend analysis and computational methods based on internal models

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

regression function and sensible statistical parameters should be demonstrated to justify that the trend actually can be used for predictive purposes.

The observation of a trend in the experimental data for a given endpoint across chemicals can be used as the basis for interpolation and may also be acceptable in certain cases for extrapolation (see Section 2.1.5 and Figure 2). Interpolation is the estimation of a value for a member using measured values from other members on "both sides" of that member within the defined category spectrum, whereas extrapolation refers to the estimation of a value for a member that is near or at the category boundary using measured values from internal category members. Interpolation can be performed when the series of values is monotonic (all increasing or decreasing) or when data are non-monotonic (e.g. parabolic). However, even in such circumstances, a substance that is not covered by other members can break the trend and show different effect. Sometimes if the level of confidence in the prediction is too low, the prediction may not be attempted.

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

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

The larger the category, the more likely that there may be breaks in trends which may affect the reliability of the interpolation or extrapolation. The observation of a "break" in a trend among some members of a category is a warning sign but is not necessarily an indication that the chemicals with different trends exhibit different toxicity pathways, but rather bioavailability of certain chemicals in the category may be affected (e.g. maximum bioaccumulation at some value of hydrophobicity and lack of other mechanisms for accumulation than passive diffusion). The bilinear or multilinear nature of trends in empirical data, if observed, can be used to confine the methods for scaling intensity of the endpoint to specific members of the category.

The observation of a trend "break" should not be confused with differences in the hazard classification of the members of a category. When the cut-off dividing different classification bands is between the extreme values of the trend, then the members of the category will be classified differently. If all members of the category have properties above or below the administrative cut-off agreed for that property, the trend analysis may be useful for judging the adequacy of forming the category but apparent breaks in the trends would not lead to differences in the classification.

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

A trend might also be expressed as a quantitative activity-activity relationship (QAAR). QAAR is a mathematical relationship between two biological endpoints, which can be in the same or different species. QAARs are based on the assumption that knowledge about the MOA, obtained for one endpoint, is applicable to the "same" endpoint in a different species, or to a similar endpoint in the same species, since

the main underlying processes are the same (e.g. partitioning, chemical reactivity, enzyme inhibition) and only the sensitivity differs. It should be noted, however, that this concept historically has been more readily applied to aquatic toxicity endpoints than for health endpoints. The use of an QAAR approach has been well established for endpoints such as estrogenicity whereby a battery of orthogonal assays has been integrated together to predict whether a substance is likely to be an agonist or antagonist (Judson et al., 2015). This network model was utilised in (Webster et al., 2019) when evaluating the WoE of a set of analogous phenols.

Thus, a chemical category can be seen as a set of “internal” (Q)SARs (and possibly also internal QAARs) for the different endpoints, with the advantage that all the underlying data are transparently available to the assessor. Such models provide quantitative descriptions of the trends within a category and are referred to as “internal” (Q)SARs (or QAARs) because these are derived directly from the experimental data for the category members. These models are also likely to be “local” models in the sense that these are based on a defined data set. Such an internal local model was developed for acute aquatic toxicity for the category of long-chain alcohols (C6- primary aliphatic alcohols) assessed within the OECD HPV Chemicals Programme OECD Cooperative Chemicals Assessment Programme.

Such methods work best for homologous series of chemicals where the metric for extrapolating from one chemical to another is a simple molecular weight, number of carbon atoms or a similar parameter which can be linked to physicochemical properties of the chemicals. However, when the members of the category are not a simple homologous series, it is essential that some parameter which predicts the trend across the members be established in order to extrapolate the measured values to the missing values. For example, the vapour pressure is mechanistically related to the acute inhalational toxicity (LC50) of ethers (Hart and Veith, 2007) because it is a surrogate for the thermodynamic activity of the chemical in the blood and tissues; but acute inhalation toxicity is not directly related to carbon number or molecular weight because the degree of branching may be significantly different among the category members. Therefore, an approach using carbon number would not produce defensible extrapolations within this category whereas vapour pressure is a more reliable parameter to extrapolate the results from measured values to missing values.

3.3.1. Examples of trend analysis and breakpoints

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

Experimental basis

A breakpoint was noted at C13-14 in the aquatic toxicity of long chain alcohols (C6-C22) as documented in the OECD HPV chemical category (OECD, 2006b). Other breakpoints have been documented, for example the sensitisation potency of cinnamic aldehydes (Patlewicz et al., 2001) as cited in (ECETOC, 2012) based on alkyl chain length. Longer chain cationic surfactants were found to exhibit reduced eye irritancy (Patlewicz and El-Deredy, 1999).

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

Computational basis

Examples on how the QSAR Toolbox can be used for filling a data gap using trend analysis and for defining an internal model can be found on the OECD Guidance Document No. 102 (OECD, 2014d). The data for a particular endpoint can be used to construct a (Q)SAR that describes the properties of the members of the category. An example of a (Q)SAR is the prediction of acute toxicity to an invertebrate species (*Tetrahymena pyriformis*) by means of a regression equation with the partitioning behaviour (*log* K_{ow} value) of the chemical as a descriptor (Schultz et al., 2002).

3.4. Computational methods based on external models

"External model" is used in distinction to the "internal model" described above and can refer to any model ((Q)SAR, QAAR, or expert system) that was not developed as part of the category formation process. Such models can be applied both in the hypothesis generation and category formation (as for example done in the QSAR Toolbox) or they can be applied to predict the read-across endpoint. If such models are used to fill data gaps in a category, these could be based on experimental data that are obtained from a wider range of chemicals than those used in the category. Such external models can be "local models" for a congeneric series of compounds which is broader than the considered category or these can be as "global models", i.e. models based on a large and diverse set of training chemicals. The validity of the "external" (Q)SARs ideally should be assessed according to five (Q)SAR validation principles and it should be assessed whether the target substance lies within the applicability domain of the model (OECD, 20014c; OECD, 2024a)). The QSAR Assessment Framework (QAF) provides support for a systematic assessment of the model and the prediction (OECD, 2024a).

It should be noted that the QSAR Toolbox profilers applied in the formation of categories and development of the read-across hypothesis, are not QSARs and have not been assessed according to the five OECD (Q)SAR validation principles (OECD, 2004f) and do not have defined applicability domains. Some expert systems apply a combined approach in which the substances might be first split using some chemical or mechanistic rationale, and models are developed for the subgroups. It should be considered, however, that it cannot be expected that (Q)SARs are available or can predict all types of substances and all types of endpoints. For complex health endpoints the read-across technique might be more informative than employment of statistical models if this is not supplemented with a detailed explanation about the predictions made. In addition to traditional (Q)SAR models underpinned by conventional *in vivo* or *in vitro* data, computational models could be underpinned with supporting data – either a structural based prediction model to predict *in vitro* HTS assay outcomes (the CERAPP and COMPARA models developed for ER and AR binding are two such examples; see (Mansouri et al., 2016; 2022)) or models based on supporting data as input to predict other *in vitro* or *in vivo* endpoints (the ER model developed by (Judson et al., 2015)) .

The predictions made by an external model may be used to provide additional support for the trend (even though reliance is usually placed on the experimental data rather than the model estimates, although predictions may also contribute to the assessment of the validity of experimental data). Especially in relation to predicting the read-across endpoint the predicted value should be compared with the experimental value available for other members of the category or an appropriate analogue. For example, a parabolic (Q)SAR could be used to characterise the trend in bioconcentration factor (BCF)[65] values across a series of substances of increasing MW.

In other cases, model predictions may be used to identify additional analogues and rationalise, per endpoint, the category members that deviate from a trend. For example, a (Q)SAR or expert system might

[65] Indicator of a chemical substance's tendency to accumulate in the living organism.

indicate that certain substances in a series have anomalous behaviour due to metabolism. Such an analysis should be confirmed by consideration of the biological plausibility of the differences.

(Q)SAR models can also provide a number of useful insights for analogue identification and evaluation; for example, the model descriptors could be used to as a basis to search for potential analogues, be used to build local QSAR models to facilitate a trend analysis within a category or similarities on the basis of the features would be used to perform a similarity weighted activity read-across prediction.

If multiple experimental data are available for a single substance, the result of a computational model can be helpful in choosing a valid data point particularly in cases whether there might be a deficiency in the test design or its reliability. SAR expert systems may also be a source of structural analogues for selected endpoints.

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

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

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

When making a prediction using a model, there are formats available providing information to facilitate regulatory consideration of both the model used and the prediction made. These formats were developed by the European Commission and are publicly available[66]. The (Q)SAR Model Reporting Format (QMRF) follows the OECD principles for the validation of (Q)SARs (OECD, 2004a). A QMRF inventory is maintained by the European Joint Research Centre (JRC) that can be utilised as a resource of QMRFs and its reference number can be referred to JRC QSAR Model Database[67]. The (Q)SAR Prediction Reporting Format (QPRF) enables the presentation of information necessary to assess robustness of the individual prediction. The newest version the QPRF format is included in the QAF guidance (OECD, 2024a), and an editable format of QPRF is available for download on the OECD webpage[68]. If using (Q)SAR data, the name, version, and owner of the models used for deriving (Q)SAR estimation data should be provided in these formats. It is recommended to use the latest software version from the manufacturer as (Q)SAR models evolve rapidly.

External (Q)SAR predictions, if valid[69], could generally be included in a WoE approach (2019a) even if experimental data are available, especially when experimental data are of limited reliability or conflicts with each other, or for difficult-to-test substances. However, this does not mean that a read-across between (Q)SAR predictions should be used for data gap filling (in case there is no experimental data). Adding

[66] EC DG JOINT RESEARCH CENTRE (2008) Institute for Health and Consumer Protection QSAR Prediction Reporting Format (QPRF) (version 1.2, September 2008).

[67] https://data.jrc.ec.europa.eu/dataset/e4ef8d13-d743-4524-a6eb-80e18b58cba4

[68] https://www.oecd.org/chemicalsafety/risk-assessment/qsar-assessment-framework-annex-2-qsar-prediction-reporting-format.docx

[69] Evaluated according to the (Q)SAR validation principles (QAF references to (Q)SAR principles (OECD, 2024a)).

(Q)SAR predictions to experimental results is particularly useful if it may help in suggesting a mode of action of the chemicals assessed.

4 Analogue Approach: A Stepwise Procedure for Identifying Analogues and Read-across

Box 4.1. Chapter 4 summary

This chapter:

- Focuses on practical aspects for forming and documenting the analogue approach.
- Provides guidance on the stepwise procedure for the analogue approach.

This guidance document can be used in a "modular" fashion and therefore making it possible to use discrete parts of the guidance by themselves. Accordingly, a number of text repetitions remain necessary in Chapters 4 and 5.

4.1. Introduction

This chapter provides guidance on how to fill data gaps as appropriate for a single or limited number of substances using the analogue approach.

The guidance in this chapter was originally based on experiences gained in the early 2000s from several regulatory programmes such as the OECD HPV programme[70], the US EPA HPV Challenge Program[71], the classification and labelling group of the EU, (ECB, 2005; Comber and Simpson, 2007; Gallegos Saliner et al., 2009; Hart, 2007; Hart and Veith, 2007; Schoeters and Verougstraete, 2007), the Existing Substances Regulation in the EU Existing Substances Programme (Tsakovska and Worth, 2007), and the Notification of New Substances (NONS) scheme (Hanway and Evans, 2000).

Read-across was extensively relied upon since 1998 as part of the OECD Cooperative Chemicals Assessment Programme. The assessments remain available through the OECD Existing Chemicals Database[72]. Since 2015, the OECD IATA Case Studies Project[73] has resulted in a number of Integrated Approaches to Testing and Assessment (IATA) cases where read-across has been undertaken.

[70] https://hpvchemicals.oecd.org/ui/Default.aspx.

[71] EPA HPV Challenge Program

[72] https://www.eumonitor.eu/9353000/1/j9vvik7m1c3gyxp/vitgbghsesz1

[73] https://www.oecd.org/en/topics/sub-issues/assessment-of-chemicals/integrated-approaches-to-testing-and-assessment.html

Read-across became the most frequently used adaptation to address information requirements under the revised EU chemicals management legislation REACH[74]. An analysis of the options used by registrants to fulfil the information requirements for the 12,439 substances registered under EU REACH, including adaptations (such as also waivers, weight of evidence and QSARs), showed that read-across was available for about 22.8% of cases in average for all tonnages (ECHA, 2017d, 2020, 2023a). Accordingly, the experiences of applying read-across under EU REACH has played a significant contribution to the updating of the guidance here.

Since the EU REACH regulation came into force in 2007, there has been a concerted increase in chemicals management legislation globally that requires the submission of test data by registrants. Notable EU-REACH like legislation exist in Korea (Chemical Substances Control Act, 2015), Türkiye (KKDIK, 2017) and the UK (UK REACH, 2021). The experience of using read-across from these registration regimes has also contributed to the guidance.

Read-across continues to play a major role in other jurisdictions, as part of Canada's Chemical Management Plan (CMP), within both the new and existing substances programmes. The Ecological Risk Classification (ERC) approach was developed by Environment and Climate Change Canada in 2016 to prioritise organic substances in the CMP and was submitted as an IATA Case Study in 2017 (OECD, 2018a). Feedback from the Case Study review helped to improve aspects of the approach incorporated in the second version (ERC2)[75]. Read-across also continues to play a major role under the amended Toxic Substances Control Act (TSCA) in the US (EPA, 2016) and within the Superfund Provisional Peer Reviewed Toxicity Values (PPRTVs) programme.

The guidance in this Chapter documents the stepwise approach of identifying and evaluating source substances as part of an analogue approach. This chapter also describes how information on physicochemical properties, chemical reactivity and, when available, metabolism, bioactivity data and MOAs should be gathered and combined with expert judgment to form a robust and well rationalised analogue approach. The choice of analogue is normally straightforward, as any potential analogue must be data-rich to form a basis for comparison. In many cases, the choice is governed by the availability and accessibility of data on an analogue manufactured by the same producer or an analogue for which data are available (e.g. the OECD Cooperative Chemicals Assessment Programme, the OECD IATA Case Studies Project, OECD Member Countries and ECHA CHEM database[76]) from the open literature.

- In the case of single substances, or complex substances (e.g. UVCBs) where there are dominating constituents, read-across often involves the identification of a chemical substructure that is common to the target substance and its analogue(s) (or their respective breakdown products) and one of the following assumptions:
- In the case of qualitative read-across, the presence (or absence) of a property/activity for the chemical of interest (target substance) can be inferred from the presence (or absence) of the same property/activity for the analogue(s) (source substance(s)).
- In the case of quantitative read-across, the known value of a property for the analogue (source substance) can be used to estimate the unknown value of the same property for the substance of interest (target substance). In the case of a toxicological effect (human health or ecotoxicological), this assumption implies that the potency of an effect shared by the two substances is similar.
- In the case of complex substances, the basis for comparison is likely to be different. For example, complex substances derived from certain process streams having similar composition may largely

[74] https://echa.europa.eu/regulations/reach/understanding-reach

[75] https://www.canada.ca/en/environment-climate-change/services/evaluating-existing-substances/science-approach-document-ecological-risk-classification-organic-substances-erc2.html

[76] https://chem.echa.europa.eu/

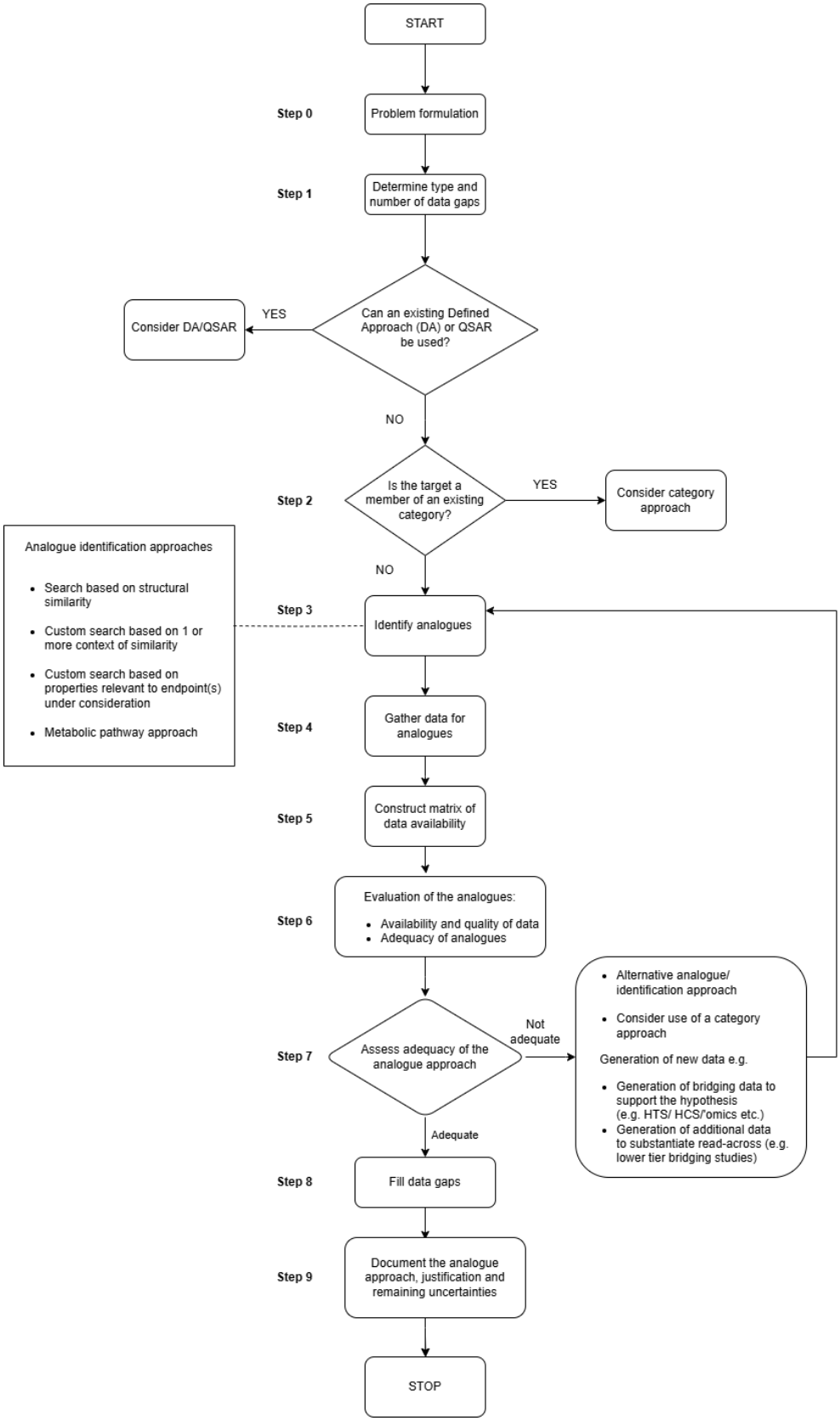
share common structures. In addition, read-across from discrete substances to UVCBs are often used when the UVCB's major constituents are known.

With limited information, it can be difficult to judge the degree of uncertainty associated with the assumption of commonality for a particular read-across. To provide the most robust read-across possible, other relevant properties should be compared between the source and target chemicals, (e.g. biological properties and bio-activation processes). The publication by (Wu et al., 2010) described a framework for identifying analogues and evaluating their suitability for filling data gaps. Other frameworks have also been published (Wang et al., 2012; Schultz et al., 2015; Patlewicz et al., 2018; Escher et al., 2019) to provide additional guiding principles related to analogue identification and evaluation as well as how supporting data can be used to inform the process. Analogues are categorised to reflect assumptions and uncertainties inherent in their use. Metabolism evaluation, various types of similarity between source and target substances, and knowledge of key biochemical processes leading to an effect (i.e. AOP) play an increasingly important role in the identification of suitable analogues and in making predictions for the target substance. This is discussed in further detail in Chapters 2 and 3.

4.2. Stepwise approach to read-across using the analogue approach

The following stepwise approach is recommended but should be regarded as flexible and not the only possible approach (see also (Patlewicz et al., 2018; Benfenati et al., 2019; Escher et al., 2019). Figure 4 provides an illustration of this approach.

Figure 4. Stepwise approach to an analogue approach



4.2.1. Step 0. Determine the problem formulation

The first step in the analogue workflow is to determine the problem formulation as this will inform the resources needed and the level of uncertainty that can be tolerated for the read-across being performed. The level of effort, the scope of data generated or collected will differ for the regulatory context or level of exposure. Understanding the decision context will enable a determination of what is sufficient for use. The problem formulation includes the purpose of the read-across, the decision context, the endpoint(s) being considered as well as the identity and characterisation of the target substance.

4.2.2. Step 1. Determine the type and number of data gaps

Use available data sources to determine what data gaps exist for the target substance. Based on the number and type of data gaps (the latter making reference to the (toxicity) endpoints under consideration, including the route of exposure), a decision can be made on whether an analogue approach is merited or whether the data gap for the endpoint of interest can be sufficiently addressed by other techniques such as (Q)SARs, IATA, or DAs, as appropriate.

For example, if the data gap was for a physicochemical property such as *log* K_{ow} or an aquatic toxicity endpoint such as acute fish toxicity, existing valid[77] (Q)SARs could be applied. Data gap filling could also be achieved using IATAs based on *in vitro* models, toxicokinetic data and models that assess relevant effects, depending on the scope and decision context. While IATAs include aspects of expert judgement, a similar approach using fixed data sources and standardised data interpretation procedures remove expert judgement. If the data gap was for an endpoint with a well-established AOP, such as is the case for skin sensitisation, then the published DAs (OECD, 2025b) which rely on assays that characterise each of the KEs could be applied.

If there are several data gaps and/or some of these are for more complex endpoints, such as repeated dose toxicity, reproductive/developmental toxicity effects, then an analogue (the subject of this chapter) or category approach (described further in Chapter 5might be the most promising strategy for data gap filling.

4.2.3. Step 2: Check whether the chemical is a member of an existing category

Having determined that a grouping and read-across approach is warranted, the next step is to establish whether the target substance is a named member of an existing established category. If it is, the category with its associated data may address the data gaps that were defined in Step 1 and a category approach should be considered (see Chapter 5). Information resources for the most common existing categories include:

- US EPA: https://comptox.epa.gov/dashboard/
- OECD Existing Chemicals Database: https://hpvchemicals.oecd.org/ui/Default.aspx
- eChemPortal: www.echemportal.org
- QSAR Toolbox: https://qsartoolbox.org/

Note that searching to determine whether a target substance is a member of an existing category is not the same as determining whether said target substance falls within the boundaries and scope of an existing category as a potential new member. This is discussed in Chapter 5, if the target chemical is not a member of an existing category, the next step is to design a strategy for identifying similar analogues.

[77] Evaluated according to the OECD (Q)SAR validation principles (QAF references to (Q)SAR principles (OECD, 2024a)).

4.2.4. Step 3: Identify analogues

There are several different ways to identify potential analogues as source substances with data with which the target substance can be compared. In cases where there is no presumption/restriction of what analogues to use, one can rely on a number of tools and techniques to assist and facilitate the identification of analogues. Tools and techniques for identifying analogues have been discussed in more detail in two ECETOC Technical Reports as well as more recent reviews (ECETOC, 2010; ECETOC, 2012; Patlewicz et al., 2017; Madden et al., 2020; Irwan et al., 2024). Some of these tools facilitate the identification of potential analogues (with and without data - which might be investigated in a subsequent step), whereas others can be searched to find associated data on a substance-by-substance basis.

The identification strategy is an exploratory process and is not intended to be an element of the read-across rationale. A systematic search strategy may identify additional potential analogues for comparison, and if a significant number of analogues are identified, then a wider category approach may be justified, as discussed in Chapter 5.

In terms of the systematic search strategy, common analogue identification approaches still rely on structural similarity or substructural assessment. It is well established now that structural similarity is only one criterion used to identify and evaluate the suitability of analogues for read-across (see Section 2.1.3). Nevertheless, structural similarity can be a pragmatic first step in identifying promising analogues that could be expected to exhibit similarity in activity e.g. (eco)toxicity, environmental fate.

The most commonly used structural similarity approach relies on characterising substances by chemical fingerprints and then performing a search to identify analogues whose pairwise similarity meets a defined threshold. The similarity threshold is a quantitative measure between 0 and 1 that summarises the commonality in structure based on the presence and absence of particular structural fragments. Each chemical is encoded into a series of “bits” that indicate presence (1) or absence (0) of specific fragments within the molecule. By far the most common similarity index that is used is the Tanimoto index (Banerjee et al., 2024), which is defined as follows.

Box 4.2. Equation 1. Tanimoto index

$$T = \frac{NAB}{(NA + NB - NAB)}$$

Where:

- *NA* is number of features (“on bits”) in structure A
- *NB* is the number of features (“on bits”) in structure B
- *NAB* is the number of features (“on bits”) common to both structure A and structure B

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

In addition, whilst Tanimoto is the most common similarity index employed, other metrics such as Cosine, Dice will give rise to different scores for the same fingerprints. For a more comprehensive review of different pairwise similarity indices see (Todeschini et al., 2012). It is worth noting that depending on the chemical representation used fingerprints versus continuous features e.g. PaDEL (Yap, 2010) or TEST descriptors[78], other metrics such as Euclidean, a generalised Jaccard will be more appropriate to use.

[78] https://www.epa.gov/comptox-tools/toxicity-estimation-software-tool-test

Differences in fingerprints (PubChem, Atom-centred fragments, ToxPrints, Extended Connectivity Fingerprints, Instem [formerly Leadscope], etc.) will also cause the similarity score to differ for the same set of chemicals using the same similarity index (Lester et al., 2018). Users should be careful using specific fingerprints from PubChem. These can have a tendency to artificially inflate or deflate similarities of substances based on small changes in the structure due to an imbalance of the chemical substructures represented. Smaller chemical structures also have a tendency to result in lower similarity scores by virtue of fewer features being represented.

Whilst there are heuristics on what might be an optimal threshold for what is highly similar based on a Tanimoto score, it is important to bear in mind that the analogue identification is an initial step – focusing on the exact threshold has the potential to exclude from consideration high-quality source analogues that may have a lower similarity or to place too much emphasis on source analogues with higher similarities that are ultimately lower quality.

Although the most common means of searching for analogues is on the basis of structural similarity using different chemical fingerprints, it is also possible to perform more targeted searches using specific features e.g. structural alerts for particular endpoints, as well as custom fingerprints that combine structural, biological and metabolic information or descriptors identified from a QSAR for a specific endpoint. Searching on the basis of structural characteristics alone is an unsupervised technique whereas a custom search, targeting features known to be pertinent for the endpoint being read across, can be termed a supervised approach. Analogues could also be searched based on bioactivity similarity alone – in this case both target and source analogues would need to have been tested in the same array of assays. Similar approaches have been successfully used in the pharmaceutical industry.

Whilst considering similarity, it is pertinent to consider the impact of any dissimilarity in the approach (structure, bioactivity etc.) and consider how these may affect the grouping and read-across strategy.

Although there are some commonalities in the contexts of similarity assessed for nanomaterials, Chapter 6.8 provides more specific details.

Tools and approaches for analogue identification.

Tools such as PubChem, freely available as a service on the National Library of Medicine website[79] provide the means to search on the basis of Tanimoto similarity or substructure across inventories of chemicals. Many of these inventories contain links to available databases or literature information. GenRA[80] provides a means to search for analogues (with associated *in vivo* toxicity data) based on different fingerprints. These fingerprints include both chemical fingerprints such as Morgan chemical fingerprints (Rodgers and Hahn, 2010), ToxPrints (Yang et al., 2015), or Analog Identification Methodology (AIM)[81] fragments (Adams et al., 2023), to ToxCast hitcall outcomes or custom hybrids of combinations that are user defined (Helman et al., 2019; Patlewicz and Shah, 2023). An example of such a search for candidate analogues might aim to optimise two different similarity contexts at the same time with a user defined weighting scheme e.g. 50% structure 50% bioactivity similarity. (Gadaleta et al., 2020) also developed a framework to search for analogues using 3 different contexts and then taking the top N analogues that overlapped more than 1 context. The approach relied on metabolic similarity (using transformation pathway similarity predicted by one metabolism prediction tool), bioactivity similarity on the basis of CHEMBL assay data as well as chemical structural similarity using typical chemical fingerprints. (Lester and Yan, 2021) applied a matched molecular pairs approach to analogue identification on the basis of structural considerations. They later extended the approach (Lester et al., 2023) to include other contexts of similarity including structural

[79] https://pubchem.ncbi.nlm.nih.gov/

[80] https://www.epa.gov/comptox-tools/generalized-read-across-genra

[81] https://www.epa.gov/tsca-screening-tools/analog-identification-methodology-aim-tool

alert profiles from tools such as Derek Nexus, physicochemical similarity and metabolic transformation information. Fingerprints representing metabolism, reactivity, and physicochemical properties were applied to 14 case study chemicals previously used to validate the framework for analogue selection by (Wu et al., 2010). To characterise similarity in metabolism, the product of the similarity scores representing the metabolic pathway of the compound and the presence of reactive species resulting from metabolic biotransformation or limited detoxification was used to heavily weight any differences in metabolism between the two substances. A fingerprint listing structural alerts for each substance was used to compute similarity in reactivity. Finally, a fingerprint for comparing the physicochemical properties for two substances made use of four properties, *log* K_{ow}, MW, charge, and volatility.

Other tools are also available including Instem (formerly Leadscope)[82] which is a commercial tool that enables sub-structural or similarity searches that can be filtered to present results for only those analogues that have associated information. The US EPA's AIM works on a different basis. Rather than using a scoring scheme such as a Tanimoto index, the set of fragments and structural features that are encoded in the programs used by the US EPA as part of its estimation Toolbox, are used as a means of identifying similar analogues with associated data. The AIM fragments have since been re-codified to facilitate analogue searching using other tools including GenRA (Adams et al., 2023; Patlewicz and Shah, 2023). There are other searching tools, and a non-exhaustive list is provided in Table 13 at the end of this chapter for illustrative purposes.

One of the most extensively used tools, in particular for EU REACH, is the OECD QSAR Toolbox (OECD, 2017g). Conveniently packaged with inventories of chemicals and several different databases, it provides a means of identifying analogues with available data from many sources. There is a plethora of methods within the QSAR Toolbox to identify analogues using one or more of the six types of profilers. These are namely predefined, empiric, general mechanistic, endpoint specific, toxicological and custom. For example, a substance could be profiled on the basis of functional groups using one of the empiric profilers or on the basis of chemical categories using one of the predefined profilers. Alternatively, a substance may be profiled by structural alerts for protein and DNA binding using the general mechanistic profilers. Searches can be tailored to retrieve analogues that might be more general in nature (e.g. structural similarity) or more specific for an endpoint of concern (e.g. mutagenicity). Extensive guidance for the functionalities and use of the QSAR Toolbox are provided here https://qsartoolbox.org/features/profiling/.

Consideration of similarity and dissimilarity for analogue selection

It is necessary to consider structural similarity as well as impacts of structural differences when selecting analogues for the read-across approach. For example, although the analogues share a common structural motif, perhaps the substituents that differ between the analogues may modulate the toxicity observed.

Overall, a hypothesis should be established, why the target and source substances are sufficiently similar to conduct a read-across approach and use the data from the source to fill the data gap for the target.

Table 11 contains examples of analogues whose structural differences were discussed in the OECD IATA Case Studies Project (OECD, 2016b).

[82] https://www.instem.com/solutions/

Table 12 shows examples of factors related to structural differences that might impact on toxicity (OECD, 2017a).

Table 11. Examples of structural differences of analogues discussed in the OECD IATA Case Study Project (OECD, 2016b)

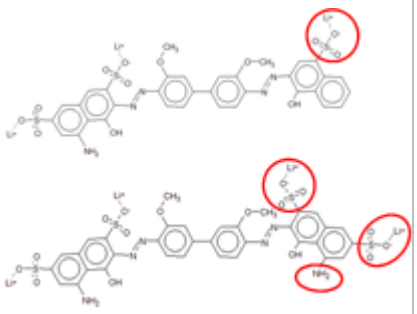
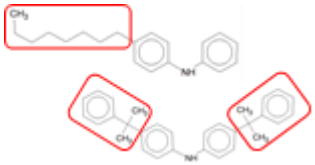
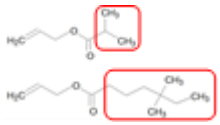
Substance /End Point	Example of analogues compared (Structural differences highlighted)	Aspects of structural differences discussed
3,3' Dimethoxybenzidine (DMOB) based direct dyes /*In vitro* mutagenicity (Ames test in *Salmonella typhimurium* under reductive conditions)		The target effect is caused by a common metabolite (released aromatic amine DMOB): extent of the influence of the number of sulfonate salt substituents in the azo substance and related solubility on this metabolic transformation
Substituted diphenylamines (SDPA) /Oral repeated dose toxicity		Subgroups were formed to better account for the differences related to type and degree of substitution: different numbers of alkyl chain substitutions or phenyl substituted derivatives
Allyl ester category /Repeated dose hepatotoxicity		The toxic effect is related to hydrolysis of the ester. There is uncertainty with respect to the range of the ester hydrolysis rate between analogues, no clear match with structural complexity, e.g. degree of branching and structural variation
4,4'-Bis (chloromethyl)-1,1'-biphenyl /Bioaccumulation		Hypothesis that the structural differences do not result in a significant effect with respect to the target endpoint

Table 12. Examples of factors related to structural differences that may impact on toxicity (OECD, 2017a)

Substance /Endpoint	Commonality in structure Variation in structure	Main factors related to the structural differences which might impact on the toxicity	Observations for toxicity effect
Phenolic benzotriazoles /Oral repeated dose toxicity	Phenolic benzotriazole	MOA: hepatotoxicity	Similarity in primary target organ: liver
	Substituents in ortho and para positions to the hydroxyl group of the phenolic ring, benzotriazole ring: no substituent or Cl at the 5 position	Substitution may influence physicochemical properties, bioavailability and toxicological properties - small structural changes result in different toxicity levels	Differences in liver toxicity levels
n-Alkanols /90-day oral repeated dose toxicity	n-Alcohol (unbranched saturated primary aliphatic alcohols)	MOA: nonpolar narcosis	Similarity in effect: no systemic toxicity; unspecific, reversible interactions with biological membranes, leading to mild effects such as increased liver weight
	Chain length (C5-C13)	Chain length affects most physicochemical properties (e.g. Low K_{ow} values increase with increasing chain length)	Similarity in toxicity levels, no trend observed in chronic toxicity. The narrow range of chain length for the applicability domain limits its impact., other than possibly on bioavailability *
2-Alkyl-1-alkanols /90-day oral repeated dose toxicity	2-Alkyl-1-alkanol (2-position branched saturated primary aliphatic alcohols; C5-C13)	MOA: nonpolar narcosis	Similarity in effect: no systemic toxicity; only mild changes: set of non-specific symptoms, including clinical symptoms, haematological values outside the normal range, or whole-body effects different from normal
	Length of the backbone (C5-C11) and length of the alkyl-substituent (C1-C3)	Changes in C-atom number affect most physicochemical properties (e.g. Low K_{ow} values increase with increasing chain lengths)	Similarity in toxicity levels, trend in physicochemical parameters not toxicologically significant for the endpoint considered and does not significantly affect bioavailability in chronic oral exposure. The narrow range of chain lengths for the applicability domain limits the impact. *
Substituted diphenylamines (SDPA) /Oral repeated dose toxicity	Diphenyl amine	MOA: induced by common functional group (diphenylamine core)	Similarity in primary target organ: liver
	Type and degree of substitution: different numbers of alkyl chain or phenyl substitutions	Type and degree of substitution and number of carbons on the side chains influence physicochemical properties and kinetics such as the oral bioavailability and metabolic pathway	Differences in liver toxicity levels Differences in target organ: spleen for monoalkylated SDPAs
Allyl ester category /Repeated dose hepatotoxicity	Allyl ester	MOA: common toxicant, metabolite allyl alcohol	Similarity in liver effect: hepatocyte and bile duct
	Chain length, degree of branching	Ester hydrolysis rate to allyl alcohol, toxicant metabolite	Differences in liver toxicity levels

* Differences in the trend of the toxicity levels between n-alkanols and 2-alkyl-1-alkanols are due to differences in the toxicokinetics profile between the two groups, e.g. differences in metabolism.

Sometimes the results of the computational search queries will identify substances which contain more than one isomer, which can give rise to difficulties in estimating the properties of the individual constituents (see example in (Worth and Patlewicz, 2007)). For example, isomers typically differ in physical properties such as melting point and boiling point. Stereochemistry might also need to be considered, depending on the endpoints under consideration e.g. estrogenic activity associated with endocrine effects. Further guidance is offered in Section 6.2.

A final consideration is to gather composition and purity and impurity information for the analogues. Generally, it is important that target and analogue substances are well characterised. For specific considerations regarding the composition of UVCBs see Section 6.6.1. Differing purity or differences in

impurities or constituents could influence the overall toxicity. In some jurisdictions, the level of the impurities, or in the case of EU CLP the presence of constituents with known hazards, might also trigger classification and labelling requirements. For example, an analogue may contain a particularly toxic impurity/constituent that is not present in the target substance, making it difficult to draw robust conclusions on the toxicity. It is therefore important that the impurities/constituents are clearly described for both the target and analogue substance and that they have similar purity profiles to allow comparison. Where purity profiles differ, it is important to describe how these differences may potentially affect the toxicity of the substance(s).

4.2.5. Step 4: Gather data for analogues

For the source analogues chosen, available data should be gathered on standard physicochemical properties, environmental fate parameter(s), ecotoxicological and toxicological effects. These data may originate from studies published in the peer-reviewed literature, databases (commercial and publicly accessible) as well as unpublished study reports. The standard information required depends on the regulatory programme but for the physicochemical parameters generally includes physical state, MW, *log* K_{ow} and other partition coefficients (e.g. Henry's Law coefficient, *log* K_{oc}), aqueous solubility, particle size, size distribution, and structure, vapour pressure, melting point, and boiling point. Since these physicochemical properties provide basic information on environmental distribution, fate and bioavailability, these can often provide supporting information for the read-across. The data gathering should include all existing relevant data and not be limited to the endpoints that are mandatory within a given regulatory programme.

Data are already available on many high-volume chemicals that have been thoroughly assessed. Information on substances assessed by the OECD is available from the OECD website[83]. One of the systems with largest database worldwide is QSAR Toolbox, containing more than three million measured data points for more than 150 thousand chemicals for various (eco)toxicological endpoints[84]. Information on chemicals can also be searched via eChemPortal[85], which provides free public access to information on properties of chemicals (i.e. physicochemical properties, toxicity, ecotoxicity and environmental fate and behaviour properties). In 2025, 35 databases participated in eChemPortal. The list of data sources participating in eChemPortal[86] is not fixed and sources are added on a regular basis. Another resource for identifying information on chemicals is the US EPA CompTox Chemicals Dashboard[87] (Williams et al., 2017). This comprises physical property, bioactivity, toxicity data and predictions across over 1 million substances. Substances can be searched on an individual basis or through using a batch search. The Dashboard provides the ability to search for literature data in addition to providing ready access to structured data from its underlying databases. It also links out to other resources such as eChemPortal. US EPA's Cheminformatics Modules[88] can also be used to search for analogues that have available data. Another publicly available resource is ECHA CHEM[89] an extensive database maintained by ECHA which includes data that companies have submitted in their EU REACH registrations. Where empirical data are not available, use of modelled and predicted data especially for standard physicochemical properties can

[83] https://hpvchemicals.oecd.org/ui/Default.aspx

[84] https://qsartoolbox.org/resources/databases/

[85] http://www.echemportal.org

[86] https://www.echemportal.org/echemportal/content/participants

[87] https://comptox.epa.gov/dashboard/;

[88] https://hcd.rtpnc.epa.gov/#/search

[89] https://chem.echa.europa.eu/

be sufficient and helpful to evaluate commonality across analogues. Commercial databases may also be adequate and should be evaluated by the user.

The level of granularity of the data being collected for analogues and how it should be summarised and documented will depend in part on the decision context and the specific regulatory needs. Reporting formats are described in Chapter 7.

4.2.6. Step 5: Construct a matrix of data availability

A matrix of data availability should be constructed for the target substance and its source analogue versus endpoints (see Chapter 7). This should start with the target substance and the analogue(s) (source substance(s)) on one axis as columns and the target endpoint(s) and all other endpoints on the other axis (as rows). If multiple analogues are identified, these should be arranged in a suitable order (e.g. according to molecular weight or *log* K_{ow}). The cells of the matrix should be populated to indicate whether empirical data are available or unavailable. If possible, the cells should also indicate the available reliable key study results (see Chapter 7).

For supporting data such as from HTS, HCS or omics, data availability (and outcomes, whether binary or continuous) may be better represented graphically for example using a heatmap, stacked bar plot, radial line graph, or multi-level donut chart to illustrate consistency of profiles between analogues. Data could be summarised or aggregated by biological targets to facilitate ease of interpretation. For the particular case of omics data, the OECD Omics Reporting Framework (OORF) provides guidance (OECD 2023a). Further information on how to construct and populate a data matrix is provided in Chapter 7, including the Chemical Grouping Application Reporting Module (CG-ARM), see Appendix.

4.2.7. Step 6: Evaluate the analogues

The next step after finding analogues and their available data is the assessment of the analogues. This considers two factors: 1) the availability and quality of their underlying empirical data; and 2) the adequacy of the source analogue itself. The robustness of the underlying data supporting the source analogue will have a bearing on the read-across predictions derived as well as the extent to which that read-across will address the problem formulation outlined in Step 0.

Availability and quality of data

Data available from relevant peer-reviewed sources such as the OECD Cooperative Chemicals Assessment Programme, the OECD IATA Case Studies Project, or from hazard and risk assessment programmes in OECD Member Countries including the EU, are often used to perform read-across.

The available experimental data should be evaluated for adequacy. In this context adequacy implies data quality, relevance, and reliability (Ingre-Khans, 2019). Section 3.1 of the "OECD Manual for the Assessment of Chemicals" (OECD, 2005a) and the "OECD Guidance Document for Describing Non-Guideline *In Vitro* Test Methods" (OECD, 2014a) provides guidance on assessing the reliability of experimental data. A scoring system, such as the Klimisch scheme (Klimisch et al., 1997), should be used by the assessor to document their judgement of the reliability of the data: a study conducted in accordance with international guidelines, such as OECD Guidelines for the Testing of Chemicals, and OECD Good Laboratory Practice (GLP) is usually considered suitable. Note the Klimisch scheme was developed for the evaluation of (eco)toxicological studies only. Other schemes for evaluating the reliability of (eco)toxicity data include Science in Risk Assessment and Policy (SciRAP)[90] (Molander et al., 2014), and Criteria for

[90] https://www.scirap.org/Page/Index/ee9102de-4b17-4c3a-86b6-e3e70d6ca3d1/evaluate-reliability-and-relevance

Reporting and Evaluating ecotoxicity Data (CRED) (Moermond et al., 2016); see also overview of data quality and systematic review references in Annex C of (OECD, 2020a).

Poor quality data for a potentially good analogue would only result in poor prediction. In addition, the information needs to be provided in sufficient detail to allow for an adequate assessment, e.g. at least a robust study summary with enough information about significantly important experimental details relating to the observations and results obtained.

Regarding relevance of the data, the available studies should be evaluated regarding their relevance related to the problem formulation and scope of the read-across. Are the available studies relevant to give evidence on the hazard in question (the type of study performed), are they accepted by the regulatory framework as relevant evidence. Or for example, is the species used for the study a resident species or otherwise not suitable to be used in the problem formulation context?

Adequacy of analogues

After assessing the quality of the available data, an assessment of the source analogue relative to the target substance is performed. This considers both the consistency and concordance of the data available for the analogue relative to the target chemical as well as across endpoints. Considerations will include:

- Data for each endpoint are expected to be consistent between target and source analogue.
- Data across endpoints are expected to be consistent for target and source analogues, e.g. shorter-term studies would be expected to be consistent with longer-term studies.
- Data for both source and target substance for endpoints for which the mode of action is likely to be similar or can cautiously be assumed to be related could be expected to be consistent based on the commonality of the MIE.

Beyond assessing the data consistency between target and source analogue, an assessment of the similarity should also be conducted. Although the source analogue may have been identified on the basis of structural similarity, evaluating its adequacy should consider other similarity contexts. Overall, a hypothesis on the similarity of target and source substances should be established and evaluated to confirm suitability of the analogues.

This will also involve consideration of the similarity of the physicochemical properties between the target and the analogue thereby providing an indication that bioavailability is likely to be comparable. The extent to which profiling outcomes from structural alert schemes are consistent between source and target chemicals may provide an indication of the known or suspected mode or mechanism of action or indeed chemical reactivity. Profiling outcomes could also take the form of assay hit-calls or potency values from HTS/HCS/omics assays from which a pairwise similarity metrics can be computed to provide quantitative measures of consistency in overall bioactivity similarity between target and source analogue. Wherever possible, the relevance of the read-across of other endpoints should be evaluated in the light of the known or suspected mode or mechanism of action. i.e. there is a tangible association between targeted biological assays profiled and the endpoint whose data gap is being filled in addition to providing a measure of overall broad bioactivity similarity. Additional considerations could consider the similarity between reactive functional groups (that will be flagged using SAR schemes) or metabolic similarity e.g. similarity in metabolic transformations, similarity in metabolites. Assessment of the robustness of the analogue approach for the particular regulatory purpose is closely related to the approach chosen for filling data gaps for any particular endpoint (i.e. analogue read-across and the use of external (Q)SARs).

There are also many tools that can assist in evaluating analogue(s) and assessing their adequacy for the read-across as described above:

- Databases with *in vivo* and *in vitro* data and other systems enabling the profiling of substances according to structure, functional groups, possible mechanisms of action. The OECD QSAR

Toolbox, Toxtree, OCHEM all provide profiling alert schemes whereas resources within the US EPA CompTox Chemicals Dashboard and GenRA can help profile substances on the basis of their empirical and predicted data including *in vitro* data.

- Expert systems such as Derek Nexus, TIMES, CATALOG, HESS[91], etc. contain many QSARs or SARs. The SARs can be used to profile candidate source analogues while the prediction models can be useful to assess the applicability of the read-across, both by predicting the missing data and comparing the experimental data available and the predictions.

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

It should be noted that evaluation of the adequacy of the analogue(s) also includes to double-check that the selection of the analogue(s) has not been subject to a bias which might impact on the read-across prediction.

4.2.8. Step 7: Assess the adequacy of the analogue approach

Aspects described Section **Error! Reference source not found.** (*Elements for a read-across justification*) need to be addressed when evaluating the adequacy of the read-across analogue approach.

The ECHA Read-Across Assessment Framework (RAAF) (ECHA, 2017b) includes crucial scientific aspects of grouping that can be helpful in evaluating a read-across approach systematically.

If after evaluating the adequacy of the analogue, based on the availability, quality, and relevance of the source data, and the suitability and relevance in terms of the problem formulation and read-across hypothesis, the read-across approach is considered to be adequate, the missing data for the target chemical(s) can be read across using the data from the source chemical(s) according to the guidance in Chapter 3 (Step 8).

If the read-across is not considered to be suitable, the following options are possible:

- It may be necessary to identify an alternative analogue – the best analogue may indeed not have the relevant experimental data, so it may be necessary to choose an analogue of lower similarity in order to obtain data. To find a more suitable analogue, the identification approach could also be adapted (see Step 3 in the workflow in 4.2.4).
- The use of a category approach could be considered particularly if a number of potential source analogues with data were initially identified.
- New data should be identified or generated in a bridging study to substantiate the similarity rationales between source and target substance further and strengthen the hypothesis. Data from HTS/HCS or omics technologies may be appropriate. In this case, the use and sufficiency of such data should be evaluated in light of the problem formulation and degree of residual uncertainty that is appropriate. The uncertainty frameworks discussed in Chapter 2 can be helpful to evaluate the sufficiency of the data and read-across for the intended purpose.
- Additional data may need to be generated using conventional toxicological approaches to fill specific data gaps to substantiate the read-across approach for the endpoint(s) considered, e.g. lower tier bridging studies to prove similar toxicity profile.

If the new data does not support the analogue approach, read-across is not possible. Therefore, there may be a need to perform standard testing according to the applicable regulatory requirements.

[91] https://www.nite.go.jp/en/chem/qsar/hess-e.html

4.2.9. Step 8: Fill the data gaps

Once the analogue approach has been determined to be adequate, data gaps should then be filled in accordance with guidance in Chapter 3 (*Techniques or methods for data gap filling*).

4.2.10. Step 9: Document the analogue approach, justification, and remaining uncertainties

The finalised analogue approach should be documented in the form of a suitable reporting format (see Chapter 7). The justification for the read-across should include an explanation of the rationale, list of endpoints being covered as well as the assessment including all relevant supporting information (see Chapter 3). The information to be read across should be included in sufficient detail. Remaining uncertainties should be transparently described (see Chapter 2).

Table 13. Selected example tools for analogue searching[92]

Tool and Website	Remarks
OECD QSAR Toolbox https://qsartoolbox.org/	Freely available, the OECD QSAR Toolbox contains tools for systematic searching of analogues and databases of experimental results, as well as methods to form chemical categories and fill data gaps by read-across, trend analysis and (Q)SARs. Downloadable
US EPA's Analogue Identification Methodology (AIM) https://www.epa.gov/tsca-screening-tools/analog-identification-methodology-aim-tool	Links to publicly available, experimental toxicity data for target chemical as well as structural analogues. A downloadable software version was released in 2012 which remains available. The AIM fragments were reproduced as far as possible in a CSRML (Yang et al., 2015) format by (Adams et al., 2023) to enable similarity searching by AIM fragments. The freely available Chemotyper tool (chemotyper.org) can be used to derive fingerprint files using the newly replicated AIM CSRML. Downloadable
Ambit http://ambit.sourceforge.net	Developed by IdeaConsult Ltd. Chemical databases and functional tools, including the LRI Ambit Read-across tool comprehensive search functionality allowing structure search, substructure search, and similarity search. Online use
PubChem https://pubchem.ncbi.nlm.nih.gov/	Publicly available database from the US National Library of Medicine (NLM). Online use
ChemSpider http://www.chemspider.com/	Database containing more than 128 million chemical structures, 277 data sources. Publicly available for online use
ChemTunes.ToxGPS® https://mn-am.com/products/chemtunestoxgps/	Commercially available chemoinformatics software.
CompTox Chemicals Dashboard https://comptox.epa.gov/dashboard	Publicly available databases and integrating diverse types of relevant domain data through a cheminformatics layer, built upon a database of curated substances linked to chemical structures. Online use
COSMOS Database https://ng.cosmosdb.eu	Freely available tool for analogues searching within chem-tox database. Online use
Derek Nexus https://www.lhasalimited.org/solutions/	Commercially available expert SAR tool for profiling candidate source analogues
Generalised Read-Across (GenRA) https://comptox.epa.gov/genra	Binary and potency-based predictions for *in vivo* toxicity data stored in ToxRefDB. Ability to predict *in vitro* ToxCast assay outcomes. Evaluation of analogues on the basis of predicted physicochemical property information. Comparison of analogues based on different fingerprint representations from chemical fingerprints such as Morgan to biological fingerprints arising from ToxCast assay hitcalls. Online use or genra-py python application available for offline programmatic batch use (https://pypi.org/project/genra/)
Hazardous Substances Database (HSDB)* https://www.nlm.nih.gov/toxnet/index.html *- Database re-organized into other products	Publicly available toxicology database on the National Library of Medicine's (NLM) Toxicology Data Network (TOXNET). Online use
Hazard Evaluation Support System Integrated Platform (HESS) https://www.nite.go.jp/en/chem/qsar/hess-e.html	Freely available and expert system containing repeated dose toxicity information to facilitate hazard assessment through the development of chemical categories. System mimics the structure/platform of the OECD QSAR Toolbox. Downloadable
Danish (Q)SAR Database and Models https://qsar.food.dtu.dk	Freely available website with predictions for 650,000 substances from hundreds of (Q)SARs and including advanced search tools, and with access to DTU-developed models for real-time predictions. Developed by Technical University of Denmark (DTU) for Danish Environmental Protection Agency (DK-EPA) with support from many stakeholders. Downloadable
SciFinder https://www.cas.org/solutions/cas-scifinder-discovery-platform/cas-scifinder	Commercially available and internet-accessible portal to extensive collection of chemical and biochemical information from scientific literature and patents. Online use

[92] Please consider confidentiality, data security, third-party access, and data ownership when using online tools.

5 Category Approach: A Stepwise Procedure for Grouping Chemicals and Read-across

Box 5.1. Chapter 5 summary

This chapter:

- Focuses on practical aspects for forming and documenting the category approach.
- Provides guidance on the stepwise procedure for the category approach.

This guidance document can be used in a "modular" fashion and therefore making it possible to use discrete parts of the guidance by themselves. Accordingly, a number of text repetitions remain necessary in Chapters 4 and 5.

5.1. Introduction

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

A number of examples both past and present have shaped the chemical category concept described in this chapter:

- More than half of the substances assessed within the OECD Cooperative Chemicals Assessment Programme and published as SIDS Initial assessment profiles have applied the chemical category approach. There are 120 categories documented within the OECD Existing Chemicals Database[93].
- A retrospective analysis of the HPV voluntary programme concluded that participating chemical manufacturers filled 55% of health and environmental effects endpoints (that could otherwise have required animal testing) by applying read-across from animal tests already conducted or proposed for analogous chemicals (Bishop et al., 2012).
- The focus of the OECD Cooperative Chemicals Assessment Programme moved from assessing the hazards of sponsored chemicals to more specialised activities in the area of chemical hazard

[93] https://hpvchemicals.oecd.org/ui/ChemGroup.aspx

assessment to include the development and application of Integrated Approaches to Testing and Assessment (IATA). The majority of IATA read-across case studies submitted to the OECD IATA Case Studies Project[94] have relied upon a category approach.

Read-across is prominent in the context of the EU REACH regulation as the most commonly used adaptation to address information requirements. An analysis of the options used by registrants to fulfil the information requirements for the 12,439 substances registered under EU REACH, including adaptations (such as also waivers, weight of evidence, and QSARs), showed that read-across was available for about 22.8% of cases in average for all tonnages (ECHA, 2017d, 2020, 2023a).

EU funded research projects notably EU-ToxRisk have investigated the development of category approaches substantiated with mechanistic data based on an established AOP. One such example included the grouping of 13 short-chain carboxylic acid analogues of valproic acid (Vrijenhoek et al., 2022).

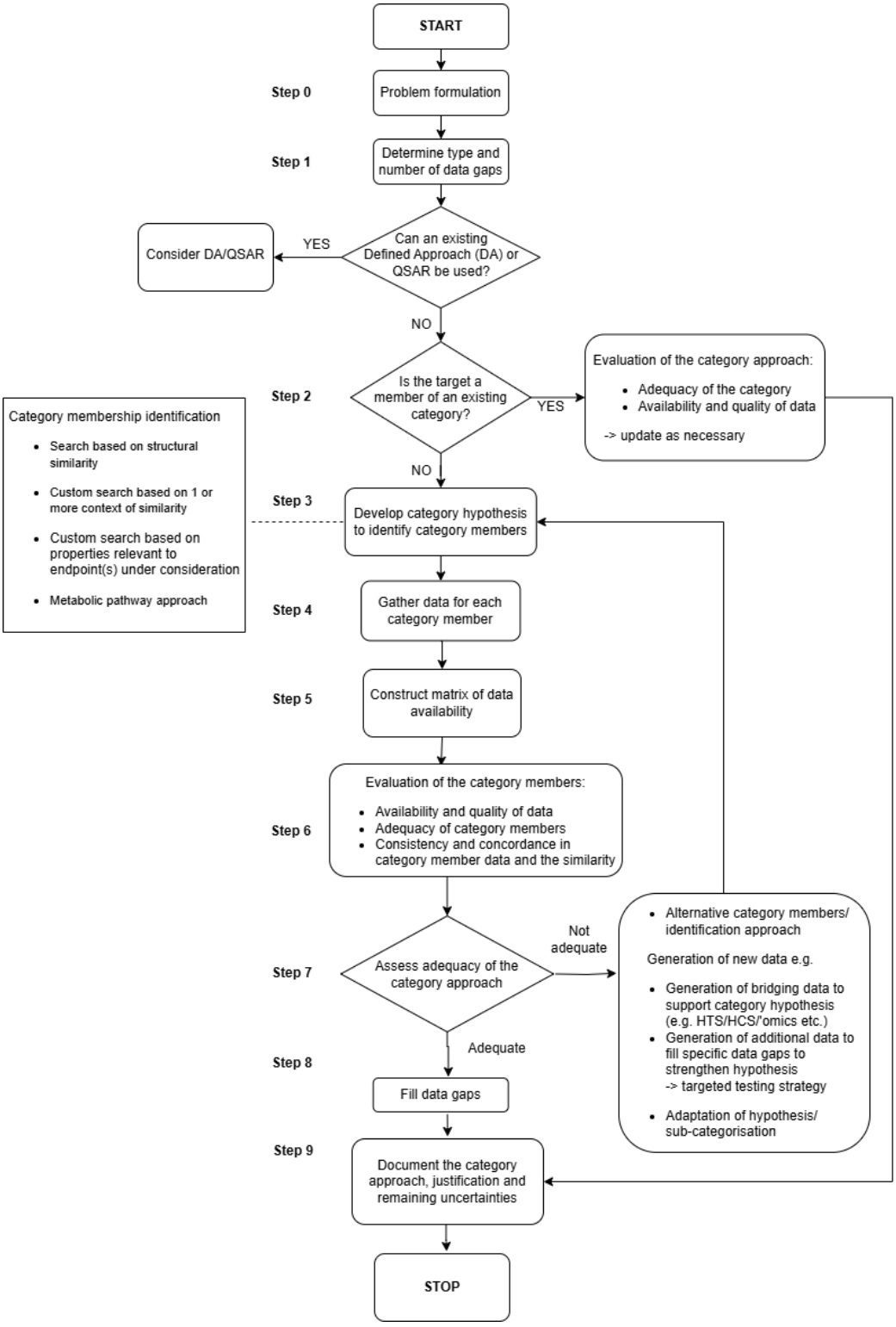
Top-down approaches of subcategorising large numbers of substances into structural groups to facilitate data gap filling efforts and/or prioritise data collection activities have been described by Date et al., (2020) for fragrance ingredients and by US EPA as part of the National PFAS Testing Strategy (Patlewicz et al., 2024a).[95]

The guidance in this chapter documents a stepwise approach to the formation of categories. This chapter should be read with the understanding that the formation of categories can be carried out using the same expertise that is routinely used in hazard identification, hazard assessment (characterisation) and risk assessment. However, given the large number and diversity of substances that exist, and the extensive number of categories that may be formed, guidance on how to develop and evaluate substance categories cannot be rigid. Rather, this section describes how information on physicochemical properties, chemical reactivity and, when available, metabolism, bioactivity data and mechanisms of action should be gathered and combined with expert judgment to form robust and well rationalised categories, as well as guidance on how to document the justification for each category.

[94] https://www.oecd.org/en/topics/sub-issues/assessment-of-chemicals/integrated-approaches-to-testing-and-assessment.html

[95] https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/national-pfas-testing-strategy

Figure 5. Stepwise approach to category development



5.2. Stepwise approach to the formation of chemical categories

In order to use the results from a category, it is necessary to demonstrate that a chemical category is both robust and scientifically justified (see Chapter 2); appropriate documentation is also essential (see Chapter 7). A stepwise approach (Figure 5) is recommended starting with a clear problem formulation and an overarching hypothesis to facilitate the identification of category members and their associated information. The general approach should be regarded as flexible, since there may be other ways of most efficiently obtaining the information needed to fill specific data gaps. Another reason for flexibility relates to the different starting points in category formation. For example, it may be possible to start from a single chemical, or small group of chemicals as seeds to build up a larger category. Alternatively, larger inventories containing relevant experimental data may be filtered down to find suitable analogues, being mindful of avoiding bias in the choice of analogues.

The steps of the workflow are also iterative, the initial hypothesis and category members might need to be adapted after evaluation of adequateness, robustness and data availability of the category developed.

5.2.1. Step 0: Determine the problem formulation

The first step in the category workflow is to determine the problem formulation as this informs the resources needed and the level of uncertainty that can be tolerated for the read-across being performed. The level of effort, the scope of data generated or collected will differ for the regulatory context or level of exposure. Understanding the decision context will enable a determination of what is sufficient for use. The problem formulation includes the purpose of the read-across, the decision context, the endpoint(s) being considered in addition to the identity and characterisation of the target substance.

5.2.2. Step 1. Determine the type and number of data gaps

Use available data sources to determine what data gaps exist for the target substance. Based on the number and type of data gaps (the latter making reference to the (toxicity) endpoints under consideration), a decision can be made on whether a category approach is merited or whether the data gap(s) for the endpoint(s) of interest might be sufficiently addressed by other techniques namely (Q)SARs, IATA, or DAs, as appropriate.

For example, if the data gap was for a physicochemical property such as *log* K_{ow} or an aquatic toxicity endpoint such as acute fish toxicity, existing valid[96] (Q)SARs could be applied. Data gap filling could also be achieved using IATAs based on a *in vitro* model, toxicokinetic data and models that assess relevant effects depending on the scope and decision context. While IATAs include aspects of expert judgement, a similar approach using fixed data sources and standardised data interpretation procedures remove expert judgement. If the data gap was for an endpoint with a well-established AOP, such as is the case for skin sensitisation, then the published DAs (OECD, 2025b) which rely on assays that characterise each of the KEs could be applied. Guidance on developing groupings based on an existing AOP are described in Sections 2.4.3 and 5.3.

If there are a number of data gaps and/or some of these are for more complex endpoints, such as repeated dose toxicity, reproductive/developmental toxicity effects, then an analogue (described further in Chapter 4) or category approach (the subject of this chapter) might be more promising strategies to address the data gaps.

[96] Evaluated according to the OECD (Q)SAR validation principles (QAF references to (Q)SAR principles (OECD, 2024a)).

5.2.3. Step 2: Check whether the chemical is a member of an existing category

Having determined that a category approach is warranted, the next step is to establish whether the target substance is a named member of an existing established category. If it is, the category with its associated data may be sufficient to address the data gaps determined in Step 1.

However, it is important to firstly verify whether the data associated with the category is sufficient to address the data gaps identified in Step 1 or needs to be updated and whether the category is in the scope of the problem formulation. If new data are available for some endpoints, these could be used to verify the robustness of the existing category and could potentially, depending on the results, lead to a revision of the category. Assessing the adequacy of the category based on the addition of new data is similar to Step 6 of the workflow, including assessing the availability and quality of data. The remaining steps of the workflow described below are then followed.

Information resources for the most common existing categories include:

- US EPA: https://comptox.epa.gov/dashboard/
- OECD Existing Chemicals Database: https://hpvchemicals.oecd.org/ui/Default.aspx
- eChemPortal: www.echemportal.org
- QSAR Toolbox: https://qsartoolbox.org/

Efforts on risk (or hazard) assessments using established categories and a category approach of substances have been made by various regulatory agencies, for example by Environment and Climate Change Canada (ECCC). Notable examples include Canadian risk assessments of aromatic azo- and benzidine-based substance grouping (ECCC, 2016). Another notable example is the structural categorisation approach developed to prioritise a large inventory of PFAS for data collection efforts by US EPA as part of its National PFAS Testing Strategy[97] (Patlewicz et al., 2024a).

A number of industry sectors have applied the principles of “grouping” for use in the assessment of health and environmental hazard properties. Examples include petroleum substances (Concawe, 2001; IPIECA 2010; House et al., 2021), dyes and pigments (ETAD, 2000), chlorinated paraffins (van Mourik et al., 2016; Kutarna et al., 2023), surfactants (CESIO, 2000; 2003) hydrocarbon solvents (HSPA, 2002), acrylate resins (UV/EB Acrylate Resins, 2003), petroleum additives (ATC, 2000a; 2000b), bitumen (Eurobitume, 2002) (ECB, 2005), and certain metals and inorganics. Categorisation approaches have also been applied to group flavours and fragrances for read-across as described in (Salvito, 2007; House et al., 2021; Date et al., 2020).

Of course, it is worth noting that querying existing categories to verify whether a target substance is a named member is not the same as determining whether said target substance falls within the scope of a category definition as a potential new member. This requires a more detailed evaluation of the adequacy and relevance of existing and proposed category members as discussed in Step 6. Determining whether a target substance may be a potential new member of an existing category relies on the category being defined explicitly with structural rules (inclusion and exclusion criteria) to enable an assessment to be made, i.e. whether the target substance falls within the applicability domain of the category. The similarity to the other category members needs to be determined to see whether the target is within the range of similarities observed across the category.

5.2.4. Step 3: Develop category hypothesis to identify category members

If alternative approaches to address data gaps are insufficient or there are no existing categories, a new category will need to be developed. The first step involves formulating a hypothesis or rationale for the

[97] https://www.epa.gov/system/files/documents/2021-10/pfas-natl-test-strategy.pdf

proposed category. This provides the practical roadmap for how category members could be identified. In many cases, this identification is conducted on the basis of a structural similarity search (e.g. phthalate esters, groups of oil-derived complex substances, metal compounds). However, members may also be identified using endpoint specific insights to tailor searches e.g. (Moustakas et al., 2022), performing custom searches addressing several similarity contexts simultaneously or by relying on a special type of category e.g. metabolic category. Chapters 2 and 3 and Section **Error! Reference source not found.** provide a more extensive discussion on the elements that can form the basis of the grouping and how these can be reported. Chapter 4 discusses the workflow for the analogue approach, of which Step 3's analogue identification, is complementary to category member identification.

Examples of chemical category hypotheses have included chemical classes with a common functional group (e.g. epoxides), chemicals with an incremental and constant change across the category (e.g. a chain length category), or chemicals having a common moiety of interest following transformation. Although chemical structure is typically the starting point in identifying members, a category could also refer to a group of chemicals related by a MOA (e.g. non-polar narcotics) or a particular property. In practice, this particular property will be largely related to the chemical structure. For example, in the case of hydrocarbon solvents, products were separated into categories based on carbon chain length, basic hydrocarbon structure – aliphatic, cyclic, or aromatic – and then further separated based on boiling ranges, carbon number, and other properties. In some cases, the aliphatic hydrocarbon categories were further subcategorised based on specific aliphatic structure such as non-branched, cyclic, or branched aliphatics (IHSC, 2004; 2005). Some categories have also been defined in terms of a metabolic pathway, i.e. these have a stepwise metabolic pathway producing the different members within the category with each metabolic step. More detailed examples of how these special types of categories have been developed and evaluated are described in Chapter 6.

Development of the category

Ideally, a category should be developed and proposed for a specific endpoint, or a selection of endpoints, rather than for all the endpoints where there are data gaps for the target substance. All endpoints that can be expected to be relevant for the category should be considered. In general, the hypothesis when starting the grouping is that all substances and all a priori defined endpoints that are carefully justified to be linked to the categorisation approach employed are covered by the category approach, i.e. the conclusions will be valid for all justified members of the category in the absence of endpoint data for some members within the category. When some members of the category present specific features (e.g. branching) known to result in different properties, for example, they may be metabolised differently, these substances will deviate from the general trend of the category. In such cases, the category approach will be limited to only those endpoints for which the data robustly demonstrate that the trend is followed, while for other endpoints, individual conclusions or conclusions for subcategories may need to be derived.

For practical or scientific reasons, a group of substances can be claimed as a category for one or several endpoints.

The category hypothesis should also address:

- The chemical similarities and trends in properties e.g. physicochemical properties and/or activities e.g. bioactivity and/or toxicity that collectively generate an association between the members. These features can be regarded as the parameters that hold the category members together.
- The MOA/mechanistic rationale, if available, that provides a basis and understanding of the read-across within the category. In some cases, this will be limited to structural alerts associated with MIEs for certain endpoints, which might give an indication of the MOA.
- The set of inclusion and/or exclusion rules that identify the ranges of values within which reliable estimations can be made for category members for the given endpoint. These rules can be described as the applicability domain for an endpoint and provide a means of potentially extending

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

It can be noted that forming a category based on lack of activity is usually subject to more uncertainties, since for example the category members cannot be linked to a mechanistic rationale or "positive" evidence connected to specific chemical reactivity, and it might be more difficult to exclude any possible effects.

The formation of a category has in many cases been dependent on which chemicals are manufactured by the consortium of companies sponsoring the category. A category may therefore contain substances that are produced by a number of different companies. Industries wishing to use a category approach should consider forming data sharing consortia (e.g. based on an industry sector group) in order to obtain appropriate support and facilitate data sharing.

There are a number of different approaches to identify potential category members from the use of simple manual approaches through to the use of computational software tools that permit searches against large chemical databases. For example, in preparing a comprehensive list of ethers to form a category of low molecular weight ethers with carbon numbers from 2 to 6, permutations of the SMILES notation for these compounds were used (Hart and Veith, 2007). This approach had the advantage of speed and simplicity, but there were also disadvantages in terms of mapping identifiers such as CAS numbers to structures and determining commercial viability. (Richard et al., 2022) developed a SMILES notation workflow for specific classes of PFAS to determine whether they met the definition of the Conference of the Parties (COP) to the Stockholm Convention on Persistent Organic Pollutants (POP) for PFOA related compounds.

Computational tools can also assist in developing the category hypothesis (rationale) in terms of its endpoints and candidate members. The choice of computational method(s) is exploratory in nature and likely to be dependent on the problem formulation.

Examples of computational tools include the QSAR Toolbox, the US EPA CompTox Chemicals Dashboard (Williams et al., 2017), Instem (formerly Leadscope), ChemTunes.ToxGPS (Altamira LLC, USA, and Molecular Networks GmbH, Germany) and GenRA (Patlewicz and Shah, 2023). A number of these analogue-searching tools are summarised in Table 31 and Table 32 in the Appendix to this guidance. Chapter 4 provides more details on the different analogue identification strategies available.

For example, the structural similarity and profiling tools within the QSAR Toolbox were used to facilitate prioritisation of substances for further regulatory action under Canada's Chemicals Management Plan (CMP).[98,99]

Sometimes the results of the computational search queries will identify substances which contain more than one isomer, which can give rise to difficulties in estimating the properties of the individual constituents (see example in (Worth and Patlewicz, 2007)). For example, isomers typically differ in physical properties such as melting point and boiling point. Stereochemistry might also need to be considered, depending on the endpoints under consideration e.g. estrogenic activity associated with endocrine effects. Further guidance is offered in Chapter 6, Section 6.2.

Although an assessment of the category members is focused on establishing similarity with respect to many contexts, e.g. structure, metabolic, bioactivity, etc., it is also important to consider the impact of any dissimilarity in the category members and how this might affect the read-across strategy performed. For

[98] https://www.canada.ca/en/health-canada/services/chemical-substances/canada-approach-chemicals/categorization-chemical-substances.html

[99] https://www.canada.ca/en/health-canada/services/chemical-substances/fact-sheets/ecological-risk-classification-organic-substances-approach.html

example, although all the category members share a common structural motif, perhaps the substituents that differ across the members may modulate the toxicity observed.

Characterisation of category members

In identifying a category, it is important that all potential category members are described as comprehensively and unambiguously as possible using relevant identifiers such as Chemical Abstract Service (CAS) numbers, structural representations e.g. SMILES, InChI, InChI Keys, and chemical names. For some substances, there may be more than one CAS number, and studies may contain relevant data reported under different CAS numbers. Due to historical reporting errors, a CAS number used to describe a substance may not accurately describe the substance as currently marketed. The CAS numbers of members of the category may benefit from being checked against different chemical inventories (e.g. Toxic Substances Control Act (TSCA), EU REACH, Customs Inventories) to identify and resolve conflicts. Confirmation from reference databases (e.g. CAS Registry, EC Number) might also be warranted. The definitive description of the substance identity will be provided from comprehensive analytical results.

Checking across these inventories can also provide an additional benefit of indicating which regulatory jurisdiction might have additional information on the substances being evaluated. For example, mapping substances present on different regulatory inventories e.g. the non-confidential TSCA inventory versus the EU REACH registered substance database have taken place as part of the DSSTox efforts within the US EPA; quality curation levels are annotated for the associations between chemical names, structures, and CAS mappings (see Grulke et al., 2019; Williams et al., 2017). These curation notes are accessible under the record information for a specific substance within the US EPA CompTox Chemicals Dashboard.

It is also important that information on the composition and the purity and impurity profiles of all potential category members is collected at the same time as details of the chemical structure. Generally, it is important that target substance and category members are well characterised. For specific considerations regarding the composition of UVCBs see Chapter 6.6.1. Differing purity or differences in impurities or constituents could influence the overall toxicity. In some jurisdictions, the level of the impurities, or in the case of EU CLP the presence of constituents with known hazards, might also trigger classification and labelling requirements. For example, a category member may contain a particularly toxic impurity/constituent that is not present in the other substances, making it difficult or impossible to draw robust conclusions on the toxicity of other category members. It is therefore important that the impurities/constituents are clearly described for all category members and that they have similar purity profiles to allow comparison. Where purity profiles differ, it is important to describe how these differences may potentially affect the toxicity of the substance(s).

At the end of the category hypothesis stage, all category members and the endpoints being covered should be listed.

5.2.5. Step 4: Gather data for each category member

For each member of the category, available data should be gathered on physicochemical properties, environmental fate parameter(s), toxicological (human health) and ecotoxicity (environmental species) effect(s). These data may originate from studies published in the peer-reviewed literature, databases (commercial or publicly accessible) as well as unpublished study reports. The standard information required depends on the regulatory programme but for the physicochemical parameters generally includes physical state, MW, *log* K_{ow} and other partition coefficients (e.g. Henry's Law coefficient, *log* K_{oc}), aqueous solubility, particle size, size distribution, and structure, vapour pressure, melting point, and boiling point. Since these physicochemical properties provide basic information to evaluate environmental distribution, fate and bioavailability, they can often provide additional supporting information for the read-across being performed. The data gathering should include all existing relevant data and not be limited to the endpoints that are mandatory within a given regulatory programme. This is particularly important to substantiate the

category justification. Where empirical data are not available, use of modelled and predicted data from (Q)SARs especially for standard physicochemical properties can be sufficient and helpful to evaluate commonality across category members. Commercial databases may also be adequate and should be evaluated by the user.

A number of the computational methods described in Step 3 (Chapter 4, Section 4.2.3, 4.2.4) can also be used to find corresponding data for the identified category members that are included in one or more databases. Having identified a range of possible chemicals, one or more databases could then be searched to identify those chemicals for which data are available. For example, category members identified within Instem (formerly Leadscope) can be then queried across other Instem (formerly Leadscope) Leadscope databases containing repeated dose toxicity or genotoxicity data. Data from various databases can be retrieved for category members identified within the QSAR Toolbox. Guidance on data gathering, with examples of databases, has already been described in Chapter 4, Section 4.2.5 as part of the analogue workflow.

The level of granularity of the data being collected for category members and how it should be summarised and documented will depend in part on the decision context and the specific regulatory needs. Reporting formats are described in Chapter 7. Specific guidance on how to prepare documentation for chemical categories with the IUCLID software are provided elsewhere (see the IUCLID User Manual, (ECHA, 2024a)).

5.2.6. Step 5: Construct a matrix of data availability

A matrix of data availability comprising category members versus endpoints should be constructed. This should start with the target substance and remaining category members on one axis (as columns) and the target endpoint(s) and all other endpoints on the other axis (as rows). The category members should be arranged in a suitable order (e.g. according to MW, or *log* K_{ow}) to more readily reflect the trends or progression observed across the category. The cells of the matrix should be populated to indicate whether empirical data are available or unavailable. If possible, the cells should also indicate the available reliable key study results (see Chapter 7).

For supporting data such as from HTS/HCS or omics, data availability (and outcomes, whether binary or continuous) may be better represented graphically for example using a heatmap, stacked bar plot, radial line graph, or multi-level donut chart to demonstrate consistency of profiles between category members. Data could be summarised or aggregated by biological targets to facilitate ease of interpretation. For the particular case of omics data, the OECD Omics Reporting Framework (OORF) provides guidance (OECD 2023a). Further information on how to construct and populate a data matrix is provided in Chapter 7, including the Chemical Grouping Application Reporting Module (CG-ARM), see Appendix.

5.2.7. Step 6: Evaluate the category members

Evaluating the category members considers two main factors: 1) the availability and quality of their underlying empirical data; and 2) the adequacy of the category members themselves. The robustness of the underlying data supporting the category members will have a bearing on any read-across predictions derived as well as the extent to which that read-across will address the needs outlined by the original problem formulation as captured in Step 0. The adequacy of the category members is complementary to the analogue evaluation step in the analogue approach already described in Chapter 4. Within the category approach, the evaluation of the members aims to address the consistency of the category members relative to the category hypothesis in terms of their similarity contexts as well as their data.

Availability and quality of data

Data from relevant peer-reviewed sources such as the OECD Cooperative Chemicals Assessment Programme, the OECD IATA Case Studies Project, or from hazard and risk assessment programmes in OECD Member Countries including the EU, are often used to perform read-across.

The available experimental data should be evaluated for adequacy. In this context adequacy implies data quality, relevance, and reliability (Ingre-Khans et al., 2019).

The available experimental data should be evaluated for adequacy. In this context adequacy implies data quality, relevance, and reliability (Ingre-Khans, 2019). Section 3.1 of the "OECD Manual for the Assessment of Chemicals" (OECD, 2005a) and the "OECD Guidance Document for Describing Non-Guideline *In Vitro* Test Methods" (OECD, 2014a) provides guidance on assessing the reliability of experimental data. A scoring system, such as the Klimisch scheme (Klimisch et al., 1997), should be used by the assessor to document their judgement of the reliability of the data: a study conducted in accordance with international guidelines, such as OECD Guidelines for the Testing of Chemicals, and OECD Good Laboratory Practice (GLP) is usually considered suitable. Note the Klimisch scheme was developed for the evaluation of (eco)toxicological studies only. Other schemes for evaluating the reliability of (eco)toxicity data include Science in Risk Assessment and Policy (SciRAP)[100] (Molander et al., 2014), and Criteria for Reporting and Evaluating ecotoxicity Data (CRED) (Moermond et al., 2016); see also overview of data quality and systematic review references in Annex C of (OECD, 2020a).

Poor quality data for a potentially good analogue would only result in poor prediction. In addition, the information needs to be provided in sufficient detail to allow for an adequate assessment, e.g. at least a robust study summary with enough information about significantly important experimental details relating to the observations and results obtained.

Regarding relevance of the data, the available studies should be evaluated regarding their relevance related to the problem formulation and scope of the read-across. Are the available studies relevant to give evidence on the hazard in question (the type of study performed), are they accepted by the regulatory framework as relevant evidence. Or for example, is the species used for the study a resident species or otherwise not suitable to be used in the problem formulation context?

Adequacy of category members

The next step after identifying category members and assessing the data quality of available data is the assessment of the adequacy of the category members. This considers both the consistency and concordance of the data available across the category members including the target substance and across endpoints.

A number of considerations will apply:

- Data for each endpoint are expected to be consistent across the category members either by demonstrating a uniform hazard or following a potency trend. Potential outliers could be indicative of breakpoints in the category and may merit subcategories to be formed to address specific endpoints.
- Data across endpoints are expected to be consistent across category members, e.g. shorter-term studies consistent with longer-term studies.
- Different types of data may be available for the same endpoint. The scope of the available results for a member of a category should not exceed the scope of the underlying data for the other members of the category, e.g. if for genotoxicity, only *in vitro* results are available for some

[100] https://www.scirap.org/Page/Index/ee9102de-4b17-4c3a-86b6-e3e70d6ca3d1/evaluate-reliability-and-relevance

members of the category (source chemicals), only conclusions on *in vitro* genotoxicity can be reached for the members of the category for which experimental results are lacking (target chemical). If the scope of the underlying experimental results for an endpoint varies (e.g. a mix of results from screening tests and higher tier tests), it is necessary to clarify the scope of the estimated results for the category members for which no experimental results are available. It may be possible to apply a WoE approach to all the data, which could lead to the same hazard identification for all the members of the category, irrespective of the data available for the individual compounds.

- An effect that is defined by a particular numerical cut-off may lead to different conclusions for individual compounds. This type of data should be studied carefully to ensure that the compounds are assessed in a way that reflects the underlying trends across a category. For instance, a series of compounds may give rise to data that shows a borderline positive irritant effect for some members of the category and a borderline negative effect for others. The data should be carefully evaluated to decide whether (a) this reflects a trend across the whole category accurately or whether (b) the uncertainties in the experimental data justify allocating the compounds to different subcategories (in this example, classifying some category members as irritants and not classifying others). If the second option is considered as the most biologically plausible explanation, the conclusion of the evaluation will lead in some cases to a different conclusion than that based on a simple evaluation of the data taken in isolation. Hence, a borderline positive effect can be interpreted as a negative effect in the light of evidence from other compounds in the category. Similarly, a borderline negative effect can be interpreted as a positive effect taking into account the data from the whole category.
- Where the data suggest possible breakpoints, the data should be evaluated to ensure that these points reflect a genuine change in properties or effects and are not due to comparison of results from testing carried out in different laboratories, at different times, with different animal strains, etc.
- The data set may contain an apparent outlier, i.e. one category member where there are experimental data that show the presence of an effect not seen in other category members. This difference can be real and provide evidence of special conditions relevant to the particular substance (e.g. the chronic and reproductive toxicity of hexane[101] compared to other lower alkanes) (Hoffman, 2008; Trimmer, 2008). Such results need to be evaluated with particular care to establish whether the result reflects a real difference in a MOA across the category or whether the test result should be questioned. Findings, which do not support the category, should be reported and interpreted to justify their exclusion since such outlying information may inform about the robustness of the category; the category might need to be reconsidered.

For some difficult-to-test and difficult-to-analyse substances (e.g., cationic surfactants), inconsistencies in (eco)toxicity and environmental fate data are expected, especially if the chemical group includes multiple substances. These inconsistencies are caused by a variety of objective reasons (mostly driven by unique physicochemical properties of these substances and by test conditions/test systems), and not necessarily by “poor grouping” and outliers.

Beyond assessing the data consistency across category members, an assessment of the similarity relative to the initial hypothesis used to identify members should also be conducted. This will evaluate the degree to which category members are structurally related (assuming that was the main identification strategy), exhibit similarity in physicochemical properties, chemical reactivity, metabolism and ADME as well as bioactivity. This assessment should be carried out for each endpoint, as the category hypothesis may lead to a relevant assessment for some endpoints and not for others and remains a matter of expert judgment. Assessment of the category rationale and robustness of the category for the particular regulatory purpose

[101] Recognised in various OECD Member Countries e.g. Japan, Korea, US, EU.

is closely related to the approach chosen for filling data gaps for any particular endpoint (i.e. read-across, trend analysis, and the use of external (Q)SARs).

For effects where the test data suggest a uniform property across a group, read-across from the existing data would normally be considered appropriate. Alternatively, the category can be sub-divided into a number of subcategories defined by the breakpoints in the category, and members assessed within each subcategory. Further subcategorisations may also be warranted to avoid high toxicity variations in very large categories.

In other cases, for example, where there is a trend in aquatic toxicity related to a change in *log* K_{ow} and based on a narcotic MOA, the data gaps may be filled by data from a valid (Q)SAR for the category.

It should be noted that evaluation of the adequacy of the category members also includes to double-check that the selection of the category members has not been subject to a bias which might impact on the read-across prediction. Adding a description of the minimum criteria used to justify inclusion/exclusion of a category member increases the transparency and is recommended, as well as tools used to support these considerations (e.g. (Q)SARs).

5.2.8. Step 7: Assess the adequacy of the category approach

If after evaluating the adequacy of the category members, based on the hypothesis, the availability, quality and relevance of the source data, the read-across approach is considered to be adequate and robust, the missing data for the target chemical(s) can be read across using the data from the source chemical(s) according to the guidance in Chapter 3 (Step 8).

If the category approach is not sufficiently robust or justified, or no adequate source data are available, one of several different options should be considered:

- Identification of alternative category members, which might entail investigation of an alternative category member identification strategy per Step 3 of the workflow (Figure 5) to find more adequate category members.
- Identification or generation of new data in a bridging study to substantiate the similarity rationales between category members further and strengthen the hypothesis. Data from HTS/HCS or omics technologies may be appropriate. The use and sufficiency of such data should be evaluated in light of the problem formulation and degree of residual uncertainty that is appropriate. The uncertainty frameworks discussed in Chapter 2 can be helpful to evaluate the sufficiency of the data and read-across for the intended purpose.
- Generation of additional data using conventional toxicological approaches to fill specific data gaps to strengthen the selected category hypothesis for the endpoint(s) considered. The filling of data gaps may generate bridging data which can support the read-across or the necessary source data. In proposing additional testing, the following factors should be taken into consideration:
  - The choice of test will be influenced by the results of the preliminary evaluation of the category (as well as any regulatory requirement).
- If there are no data for any of the members of a category for a particular endpoint, testing of a limited number of carefully selected category members may be considered appropriate; when data are already available indicating the presence or absence of a particular effect, tests may be chosen to provide evidence that compounds selected for testing show the effects that have been predicted based on the available data for the property. For example, for a substance in a category where skin irritation is predicted, a simple *in vitro* test might be adequate for hazard identification and follow-up classification and labelling and risk assessment.
- For testing of selected category members, a bespoke testing strategy should be developed and explained/justified. It should include a category definition, rationale, and matrix of data availability.

  - The testing strategy may also take account of the needs of subcategorisation where this may be expected (e.g. expectation of break points in a trend due to differences in bioavailability).
- Adaptation of the hypothesis: If further examination of the data suggests that there are similar effects or a pattern of effects only for a limited number of chemicals in the group, then the analysis might suggest that the category should be modified e.g. divided into subcategories or a chemical should be removed from the category. The hypothesis might need to be adapted (return to Step 3).

If adequate data do not exist for the/all category members, but the structure-based category is reliable for one or more endpoints, then a category approach may still be proposed for these endpoints. Testing of some chemical category members would still be necessary (see above). The choice of chemicals and endpoints for testing should be scientifically motivated but is also likely to involve animal welfare and financial considerations, especially in the case of more “expensive” endpoints.

Testing for category members: Testing strategy

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 (where justified) extrapolation to other members of the category that have no data or proposed testing. The category rationale should be documented, as described in the category reporting format in Chapter 7.

The data matrix summarises the existing data and is an important indicator of how the proposed testing will adequately characterise the category. Each endpoint should have a row in the matrix. If toxicity is expected to vary in a regular pattern from one end of the range of category members to the other end (e.g. high toxicity to low toxicity) and this trend will form the basis for predicting the hazard, data should be available (or need to be generated) for substances to bracket both ends of toxicity. If the category is large, testing also needs to be performed and/or data should be available for one or more member(s) in the middle of the range of toxicity to confirm the trend holds and to check, for example, for the occurrence of potential break point in the trend. Any change in a tendency for a property should be accompanied by data in the adjacent cells in order to define the limits for the resulting subsets of the category or subcategories. Data on rate and extent of metabolism as well as metabolite identification may need to be available if this is part of the basis for read-across. The overall information may need to address the impact on prediction of the allowed structural differences in the category. For example, the impact on the prediction of simple anions in the case of read-across based on bioavailability of metal cations.

Assuming the columns are the category members, there are no rules for the number of columns and cells that must be filled nor the number that can be empty as cases can vary in their complexity and may need, for example, category- or substance-specific considerations or depending on the certainty required for the given decision context.

When selecting a sample to test, it should be representative of the substance manufactured or imported, including the presence of any manufacturing impurities. It should also be noted that the category test plan is intended to provide information about the properties of the group as a whole rather than the properties of any specific, individual compound. A category test plan may thus identify key substances for testing that are of little or no commercial importance. While in some cases this may even require the synthesis of chemicals specifically for this purpose, the approach may still prove more economical, both in terms of expense and numbers of animals used for testing, than a more conventional testing strategy based on individual commercially available chemicals e.g. in order to cover the category comprehensively and support the category hypothesis strategically.

Ideally the substances defining the borders of the category should be included in the testing strategy, if testing has not already been performed.

Re-evaluating the category following new data generation

When new information becomes available or the category members have been updated or subcategories have been formed, the (sub-)category approach should be re-evaluated to determine whether the criteria outlined above in Step 7 are satisfied and therefore whether the category can be finalised and documented. If the results support the category, the testing phase is complete, the chemical category can be finalised and documented (Step 9). Remaining data gaps can be filled according to the guidance in Chapter 3 (see Step 8, Section 5.2.9).

If the results do not support the category hypothesis it has not succeeded in its current form and needs to be re-evaluated. For example, the new data may show it is possible to refine the scope of the original category for one or other endpoint (e.g. dividing the category as appropriate) and may also therefore allow identification of further targeted testing e.g. to strengthen confidence in the trend or confirm the boundaries of the new (sub)category if not already available. New toxicities may be identified which require further investigation.

Alternatively, the category proposal may be dropped altogether. In the latter case, there may be a need to perform standard tests according to the applicable regulatory requirements.

5.2.9. Step 8: Fill the data gaps

Once the category approach has been determined to be adequate, the considered data gaps should then be filled in accordance with the techniques described in Chapter 3 (*Techniques or methods for data gap filling*). Note if subcategories have been developed or if the category hypothesis specifies which endpoints are in scope, any data gap filling should be performed for the relevant category members only.

5.2.10. Step 9: Document the category approach, justification, and remaining uncertainties

If the hypothesis has been confirmed in the previous steps to be adequate and the category to be robust, the finalised category should be documented in the form of a suitable reporting format (see Chapter 7). The justification for the read-across should include an explanation of the category rationale - describing the common features of the category members, the reasoning for assessing the chemicals together as a group, the list of endpoints being covered as well as the assessment including all relevant supporting information (see Chapter 3). The information to be read across should be included in sufficient detail. Remaining uncertainties should be transparently described (see Chapter 2).

5.3. Category development using an AOP

This section discusses the general consideration on developing a category approach using an AOP. A full AOP from the MIE and KEs through to the final adverse outcome is not required before being able to build a chemical category around a common MOA or KE. Mechanistic information, such as that obtained from HTS/HCS or omics technologies, can be used to justify the grouping of chemicals around a given adverse outcome, provided a link can be established between the endpoint(s) in the molecular screening and the adverse outcome (see also Section 2.4.3). Positive and negative controls chemicals should also be included to further support the mechanistically based grouping hypothesis (see also Appendix (GC-ARM Annex)).

Table 14 illustrates conceptually how mechanistic information may be relevant in forming chemical categories. The table shows the various situations/scenarios when using mechanistic information in the form of profilers to justify the grouping of chemicals for a given apical endpoint. The full AOP description represents scenario 1; other scenarios are variations of available data along the pathway.

Table 14. Qualitative use of mechanistic information in forming chemical categories

Scenario	MIE	KE1*	KE2	KEn	AO**	Application/Usefulness
1	●	●	●	●	●	Most chronic effects
2	●	○	○	○	●	Acute and some local effects
3	●	●	○	○	●	Many local effects
4	○	●	○	○	●	Some use***
5	○	●	●	○	●	Some use****
6	●	●	○	○	○	No use
7	●	●	●	○	○	No use
8	○	●	○	○	○	No use

● information available, ○ information not available

*: KE is defined as key event

**: AO is defined as adverse outcome

***: some use when the KE is e.g. positive receptor binding, and the AO is reproductive toxicity

****: some use when the KE1 is e.g. receptor binding, the KE2 is vitellogenin increase or decrease in fish, and the AO is reproductive toxicity

Depending on the amount and the distribution of information/data (e.g. for several chemicals for an assay versus a few data points for an assay, or for several events along the pathway versus on a single event) utility may vary to justify the chemical grouping for a particular use. For every event listed in Table 14 (e.g. KE1), there is the possibility of having data from one or several (typically less than 10) protocols or methods assessing that event. Having data for the same chemical evaluated in different assays allows for an evaluation of the reproducibility of the event, that is of value in assessing confidence in a particular result. Conversely, having many rather than a few chemicals tested in a particular assay for an event is of greater value in assessing the confidence in the assay and the result. It is also more valuable for the category justification to have results from assays representing several different events than from several assays representing a single event. However, relying solely on a key event that is at a high level of biological organisation (i.e. a more integrative KE) is likely to run the risk of mixing chemicals where different mechanisms lead to the same apical outcomes.

Furthermore, any information on chemicals in the group formed, which are shown to trigger the MIE, or KE(s) along the AOP, will contribute to the justification of the category. For example a (Q)SAR may be used to predict the MIE (e.g. protein binding, ER binding), an *in vitro*/*ex vivo* assay may be used to support a MIE or a KE in an AOP (e.g. vitellogenin induction in fish liver slice; *in vitro* alterations in sodium flux through voltage-gated sodium channels, leading to neurotoxicity), *in vivo* data may support a key event specified in the AOP (e.g. similar specific histopathological findings or triggering of MOA related response such as vitellogenin (egg yolk protein) in blood plasma/plasma of male fish, change in sex of fish or, organ weight changes in rat, specific protein expression, specific animal behaviour).

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

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

6 Guidance on Specific Types of Categories

Box 6.1. Chapter 6 summary

This chapter:

- Elaborates on some of the specific aspects that need to be considered for grouping of chemicals, related to specific chemistry issues such as chain length, isomerism, or metabolism.
- Elaborates on some of the specific aspects that need to be considered for grouping certain types of chemical substances and forms such as ionisable compounds, substances of unknown or variable composition, complex reaction products or biological material (UVCBs), metals and inorganic compounds, and nanomaterials.

  It should be noted that different regulatory frameworks might have different requirements for specific types of substances.

6.1. Chain length

Chain length categories show an incremental, and usually constant, increase in chain length across the category. Examples of chain length categories which have been assessed within the OECD Cooperative Chemicals Assessment Programme include alpha-olefins, higher olefins or monoethylene glycol ethers[102]. In the alpha-olefins case, each category member differed by a methylene group (–CH_2– unit), whereas in the ethylene glycols category, there was an incremental increase in the number of CH_2CH_2O groups.

Categories defined by chain length generally show an incremental change in MW and other physicochemical properties, such as water solubility or *log* K_{ow}. However, not all properties will necessarily exhibit a linear relationship or even monotonous trends with chain length and care must be taken when making assumptions about the effect of chain length. For many homologous series, an increasing chain length leads to increasing *log* K_{ow} and concurrent decrease in water solubility. Depending on the MOA, this may lead to predictable changes in biodegradability, toxicity, and bioaccumulation. A decrease in water solubility and thus bioavailability may lead to decreases in biodegradability. An increase in lipophilicity may lead to an increase in bioaccumulation up to a certain *log* K_{ow}. For chemicals with a non-polar narcosis MOA, decreasing water solubility would be associated with decreasing aquatic fish toxicity. For example, in the alpha-olefins category, there was an apparent cut-off point between the C8 and C10 chain length at which acute toxicity to fish was no longer observed. Similarly, a trend of increasing MW may lead to

[102] https://hpvchemicals.oecd.org/ui/ChemGroup.aspx

decreasing systemic toxicity as absorption decreases. There may also be a change in physical state of the category members as chain length increases e.g. from liquid to solid.

Care should also be taken when evaluating a category containing both branched chain chemicals and linear chain chemicals. Branching may be a more significant driver for certain endpoints, such as biodegradation and teratogenicity. For these endpoints where differences in trend are seen, it may be helpful to divide the category into subcategories (e.g. endpoint-specific) in order to provide a robust justification for the assessment.

Careful thought should be given to selecting the boundaries of a chain length category, for example, surface activity should be considered. The cut-off points discussed already may provide useful boundaries. The potential scope and size of a chain length category may be larger than that covered by a particular manufacturer or consortium. Where possible, well-characterised substances which are not necessarily subject to a particular regulatory programme but which could fit into the series should ideally be included though it is recognised that a particular manufacturer or consortium may define a category based on the substances within their own purview. There may be cases when testing the end members of a chain length category is not appropriate. For example, if the existing data indicates that the toxicity cut-off occurs earlier in the series, it may not be necessary to test the end member for that endpoint.

(Q)SARs can be used to provide information for category members (as discussed in Section 2.1.4). In general, substances at either end of a chain length category should have all endpoints fulfilled, preferably with test data, if breakpoints have not been indicated. This would enable an interpolation of data for the other category members. For example, in the category on ethylene glycols, a linear regression was used to predict acute aquatic toxicity, indicating that toxicity decreased with increasing chain length, and further supporting the low toxicity of the category members concluded from available experimental data (OECD, 2004b). For categories where there is more than one variable, such as variation in the length and degree of branching of the chains, more category members are likely to be required to provide confidence in the inferences being made.

Other examples of chain length characteristics are oleochemical derivatives, which can be grouped in such categories as fatty acids or alkyl sulphates (OECD, 2007a). These categories may contain single-chain chemicals as well as complex substances containing chemicals of distinct chain lengths at varying amounts, with some chain lengths maybe more prevalent than others in a given sample. The proportions of different individual chain length molecules in complex substances typically mirror the distribution of chain lengths found in the natural fats and oils from which these substances are derived. Since the category chemicals differ from each other only by the number of $-CH_2-CH_2-$ units, these categories are very often homogenous and exhibit a constant pattern in the changing of the potency of the properties across the category. However, great care should be given for the fact that the functional group introduced in the natural chemicals may change the metabolism or the MOA.

6.2. Isomers

Isomers are chemicals that have identical chemical (or empirical) formula but different molecular arrangements. Although there are several types of isomers, the two that typically will be considered are structural and geometric.

- Structural isomers are molecules with differences in the arrangement of their atoms. Structural isomers can include:
  - Chain isomers. For example, hydrocarbon chains with identical or variable lengths and variable branching patterns (see also Section 6.1).
  - Positional isomers. For example, hydrocarbon chains with a functional group that varies in position along the chain (e.g. 1-butene and 2-butene).

  - Functional group isomers. These isomers also have identical molecular formulas but contain different functional groups. Examples are butanal and butanone which both have the chemical (or empirical) formula $C_4H_{10}O$. Each of these isomers contains a carbonyl group (C=O) but are representative of two different chemical families: butanal is an aldehyde whereas butanone is a ketone. This type of structural isomers is less likely to be considered within a category because functional group isomers can have very different chemical and biological properties. Functional group isomers are not included within the scope of this guidance.
  - Tautomers. These are structural isomers that have several possible arrangements of double bonds in a dynamic equilibrium. These arrangements may possess different biological and toxicological properties. For example, keto-enol tautomerism in phenolic structures (e.g. hydroquinone) can change an apparent di-phenol to a quinone that is a more activated species of the same chemical structure. Other examples are enamine-imine, nitroso-oxime, amide-imidic acid or lactam-lactim (cyclic) tautomerism. In order to predict reliably the behaviour of such chemicals, all tautomeric forms of the chemicals must be evaluated.
- Stereoisomers are isomeric molecules whose atomic connectivity is the same but whose atomic arrangement in space is different. The following stereoisomerism can be distinguished:
  - Diastereomers are non-superimposable stereoisomers (geometric isomers): these are non-mirror images of each other. Cis-diastereomers have substituent groups projecting in the same direction; trans-diastereomers have substituents oriented in opposing directions. These diastereomers can occur when a double bond or a ring is present which restrict the rotation. For example, cis-2-butene and trans-2-butene each have carbon groups on either side of a double bond, which cannot rotate, so the carbon groups are arranged on either the same side of the molecule (cis) or opposite sides of the molecule (trans).
  - Enantiomers are two stereoisomers that are related to each other by a reflection: these are mirror images of each other. Every stereocentre in one has the opposite configuration in the other. These are typically denoted R- and S- depending on their arrangement in space. S means that the configuration of the substituents from the stereocentre (going from highest to lowest priority as defined by MW) are arranged so that they go counterclockwise (or left) and R clockwise (or right). Two compounds that are enantiomers of each other have the same physical properties, except for the direction in which these rotate polarised light and how these interact with different optical isomers of other compounds, and how these interact with enzymes. In nature, only one enantiomer of most chiral biological compounds, such as amino acids, is present. As a result, different enantiomers of a compound may have substantially different biological effects. A well-known example is thalidomide, where the R-enantiomer has sedative effects, however, the S-enantiomer is teratogenic.

Stereoisomers can have similar or different chemical or toxicological properties. Even though these may behave identically in many chemical reactions, it is, for example, well known that the enzyme specificity in biological systems may be totally different, so caution is needed in case of such substances. Several illustrations of the impact of chirality on the toxicity and fate are given by (Smith, 2009).

The substance(s) with a data gap as well as substance(s) with data are similar such that their physicochemical, biological, and toxicological properties would be expected to behave in a predictably similar manner or logically progress across a defined range. The incremental change is so small that it is not expected to affect the property sufficiently. This similar manner or logical progress should be demonstrated by the available experimental data. (Q)SAR models and trend analysis can also be used in addition of experimental data to support the estimate. Extended Simplified Molecular Input Line Entry System (XSMILES) e.g. ChemAxon Extended SMILES (CXSMILES) or InChI now exist which facilitate cheminformatic analysis of isomers.

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

Based on the category of Butenes (OECD, 2004c) and their mixtures, the following observations were made:

- Selected properties of isomers may be read across to other isomer(s) or to an isomeric mixture within a category if the data are similar and/or if the structure of the isomer(s) without data is similar to the isomers with data.
- Extrapolating properties to isomeric mixtures should take into account mode of action, potential additivity and synergy, as well as purity profiles, and mixture composition.
- For toxicological endpoints (e.g. LC_{50}, NOAEL), a range of toxicity or the lowest value in a range of toxicity may be used for read-across.
- Read-across from one isomer to another may not be straightforward. Metabolic data may be needed if existing knowledge of category members or related non category members suggests that differences may be expressed within a biological endpoint of interest.

6.3. Ionisable compounds

For categories that include ionisable substances, several properties related to the behaviour of the ionisable substances need to be considered, including:

- Water solubility and dissociation behaviour in aqueous solution should be similar or follow a predictable trend amongst category members.
- pH and redox potential (Eh) effect on behaviour in aqueous solution and toxicological properties e.g. irritation
- Effect(s) of the counter-ion. It is possible that the counter-ion(s) may pose hazards of greater concern than the common cation or anion on which the category is based e.g. metal counter-ions that are inherently hazardous on their own. Under such circumstances, it may be of limited utility to group and assess substances by the component which is expected to have the least effect.
- Differentiation between toxicity stemming from pH effects alone and a combination of pH effects and inherent toxicity. Comparison of unbuffered test results with test results that include pH is preferable.
- For substances in which the anion is associated with toxic effects (e.g. cyanides, oxalates, thiocyanates), the toxicity of the anion should be considered when performing grouping.

One example is the Methanolates category assessed under the OECD Cooperative Chemicals Assessment Programme[103]. This consists of 17 potassium and sodium methanolate which both react rapidly in water to form the corresponding hydroxide. The sodium and potassium hydroxides were determined to have a low hazard for the environment however the degradation product methanol was considered a concern for human health.

When comparing acids, their salts and multiprotic molecules, differences arising from pH effects should be considered (Caley et al., 2007). For example, skin and eye irritation are likely to be different for an acid compared with its salt. This is illustrated by the Phosphonic Acid Compound (Groups 1, 2, 3) categories assessed under the OECD Cooperative Chemicals Assessment Programme. For these categories, dermal and irritation studies are considered separately for the acid and salts.

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

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

6.4. Metabolic or degradation pathways and toxicokinetics

6.4.1. Concepts

One of the rationales that can underpin a category is the occurrence of a common precursor and/or breakdown product that results via physical or biological processes (i.e. metabolic or degradation pathway similarity).

Hypotheses that underpin a metabolic/degradation pathway similarity could include:

- Substance A is metabolised/degraded to substances B, C, D etc (where A is the target and B, C or D are the source substances or vice versa).
- Substance A and B are both metabolised/degraded rapidly into substance C (where A is the target and B is the source substance or vice versa) and evidence suggests that C causes an adverse effect.
- Substance A is metabolised/degraded to substance C and substance B is metabolised/degraded to substance D (where A is the target and B is the source or vice versa) and evidence suggests that C and D cause a similar adverse effect.

The metabolites that are most often considered are downstream blood metabolites though the discussion here may equally be applicable to other organ and tissue specific metabolites.

The metabolic pathway approach considers the potential of using data from a parent chemical to characterise the hazards of the metabolites, or vice versa. Hazard identification studies with the parent chemical could be potentially used to identify the hazards associated with systemic blood levels of the downstream primary and secondary metabolites which could be used in lieu of studies conducted on the primary and secondary metabolites themselves. If the metabolism of the parent chemical within barrier tissue (e.g. lung, gut, placenta tissue) occurs so rapidly that the initial primary metabolite is the predominant

[103] https://hpvchemicals.oecd.org/ui/Default.aspx

chemical found within the blood, data from studies conducted with the primary metabolite itself can be used to characterise the hazards of the parent chemical.

Examples of metabolic analogue/category approaches have been published. *In vivo* evidence for the metabolic hydrolysis of methyl acetate to methanol and acetic acid in rabbits and humans formed the basis for justifying the read-across between methanol and methyl acetate as part of an US EPA Provisional Peer Reviewed Toxicity Values (PPRTV) assessment. Under the OECD Screening Information Data Sets (SIDS), the Acid Chlorides category was structured on the basis of a number of analogue pairs, each comprising an acyl chloride and its corresponding acid, with the rationalisation that the hazard profile of the parent acyl chloride could be assessed by reference to its transformation product. Other examples under the OECD Cooperative Assessment Programme included n-butyric acid where n-butyl acetate was used as the metabolite precursor to address repeated dose toxicity data gaps. In (Ball et al., 2014), dipropylene glycol methyl ether (DGME) and its acetate (DGMEA) formed the metabolic pair. In OECD Case Study 2018-1 (OECD, 2019b), metabolic products played a key role in the assessment of testicular toxicity of ethylene glycol methyl ether (EGME) related chemicals (OECD, 2019b).

6.4.2. Opportunities to increase the availability of experimental metabolism data

One of the first issues to consider when forming a metabolic pathway analogue or category approach is to determine what metabolism information is available or could be generated under the regulatory programme for which the chemical is being assessed. ADME/TK (Absorption, Distribution, Metabolism, Excretion/Toxicokinetic) studies are not often requested in many regulatory programmes and therefore would require additional work. *In vitro* metabolic studies in conjunction with physiologically based kinetic (PBK) or physiologically based dynamic (PBD) models (the latter if feasible) may prove sufficient to characterise the metabolic pathway(s). The OECD published Guidance on the Characterisation, Validation, and Reporting of PBK Models, especially in the regulatory use when no *in vivo* kinetic data for model validation are available (OECD, 2021a).

OECD IATA Case Study 2019-5 (OECD, 2020n) characterised the target chemical hazard using a read-across approach to other branched carboxylic acids. The case study used a rat PBK approach that was developed from *in vivo* data of one analogue in order to guide the selection of a concentration range for *in vitro* testing. A human *in vitro* to *in vivo* extrapolation (IVIVE)-PBK model (scale up) was developed based on physiochemical properties and *in vitro* clearance data (e.g. plasma protein binding (ppb) and intrinsic hepatic clearance (CLint, Hep), and combined with the most sensitive *in vitro* endpoint to derive an oral equivalent dose. The extrapolated values from the *in vitro* studies were then compared with the *in vivo* data from two analogues and were found to be in the same range as the corresponding *in vivo* animal studies. It was concluded that the target chemical did not induce hepatic steatosis up to the highest dose tested *in vitro* or the derived oral equivalent dose.

OECD IATA Case Study 2021-5 (OECD, 2022) sought to characterise the Developmental Neurotoxicity (DNT) properties of a target chemical and its metabolite. The main use of the PBK models developed was to assess the mother-plasma and foetal-brain concentrations following ingestion of single or multiple doses of the compounds, and to predict comparative human exposure of the compounds using IVIVE approaches. To help with the parameterisation of the PBK models in human hepatocytes, binding to human plasma and blood was measured *in vitro*. PBK models were then constructed using full body models in both rat and human where each organ of the body was modelled as a separate compartment. Oral absorption was predicted using the Advanced, Dissolution, Absorption and Metabolism (ADAM) model within the Simcyp simulator. The IVIVE model concluded that a brain concentration of 2 µM imidacloprid, which was considered a PoD from *in vitro* studies, would be reached after an intake of 0.2 mg/kg body weight in the average population. The case study authors recommended Table 3.1 PBK Model Reporting Template within the OECD PBK document (OECD, 2021a) as a format.

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

Alternatively, *in vitro* metabolic data which can be performed at lower cost and higher throughput than animal studies could be generated. There are two main limitations to take into account in this case. Firstly, *in vitro* studies will not necessarily reproduce all *in vivo* metabolic pathways and hence metabolites formed. Secondly, the concentrations of the chemical and its metabolites *in vitro* may differ from what might occur *in vivo*. That said, it may be sufficient to use inherent clearance rate information derived from *in vitro* studies using metabolically active cells (such as primary hepatocytes) for a well-established transformation, such as hydrolysis, to substantiate a read-across hypothesis. Indeed, a number of the cited examples above were focused on the hydrolysis of an ester or acid halide.

MetaPath[104] is a freely available software platform developed for the purpose of archiving, sharing, and analysing experimental data on metabolism, metabolic pathways, and supporting metadata. It facilitates the compilation and organisation of metabolism study results into a systematic database, enabling data comparisons and evaluations. MetaPath was originally developed by the Laboratory of Mathematical Chemistry (LMC) in collaboration with US EPA, and further advanced to facilitate the assessment of pesticide metabolites in collaboration with the European Food Safety Authority (EFSA). Currently, efforts are underway through an OECD working group, MetaPath Users Group, to extend its use, to further improve the connectivity of MetaPath with the QSAR Toolbox and IUCLID, and to facilitate the exchange of datasets between regulatory agencies.

6.4.3. Opportunities to increase the availability of experimental metabolism data using non-targeted and/or metabolomics analytical methods

Traditionally, targeted analytical chemistry methods have been needed to quantify a parent substance and potentially its metabolites (which can also be referred to as biotransformation products). However, for the last decade or more, non-targeted analysis (NTA) methods (also known as untargeted analysis) have greatly improved the ability to measure multiple parent chemicals and their metabolites in various sample types (Schymanski et al., 2015). Recently, similar analytical approaches have been applied to characterise the metabolites formed from deliberate chemical exposures within *in vitro* and *in vivo* toxicology studies in concert with chemical screening lists derived from predicted metabolites generated by *in silico* metabolism tools (Sobus et al., 2018, see Section 6.4.4), providing a powerful means to start to characterise likely metabolic pathways. For example, (Boyce et al., 2023) reported a proof-of-concept study where a combination of NTA and *in silico* approaches were investigated. Specifically, chemicals were first incubated in primary liver hepatocytes, thereafter the supernatant and lysate fractions were analysed with high-resolution liquid chromatography mass spectrometry (LC-MS). (Bowen et al., 2023) demonstrated how a LC-MS (Liquid Chromatography-Mass Spectrometry) metabolomics assay primarily applied to characterise endogenous metabolic responses to chemical exposure, both *in vivo* and *in vitro*, could also detect parent chemicals and their metabolites (or biotransformation products). This approach provides an

[104] https://oasis-lmc.org/products/software/metapath.aspx

opportunity to simultaneously gain deep insights into chemical fate and metabolism, and to associate the internal relative dose directly with endogenous metabolic effects. Furthermore, (Bowen et al., 2023) demonstrated how by applying an untargeted analytical and computational workflow to rat and human induced pluripotent stem cell cardiomyocytes, the metabolic competency of the *in vitro* model could be determined empirically relative to measurements of heart tissue from the *in vivo* model.

6.4.4. Opportunities to utilise predicted metabolism data for grouping

There are a number *in silico* metabolism tools, both commercially and freely available that can be used to generate metabolites and potentially infer metabolic transformation pathways. Examples of tools include the metabolism simulators within the QSAR Toolbox that are part of the OASIS Suite developed and maintained by Laboratory of Mathematical Chemistry (Mekenyan et al., 2004), the Nexus suite comprising Meteor Nexus from Lhasa Ltd as well as open-source tools such as Biotransformer (Djoumbou-Feunang et al., 2019) and SyGMa (Ridder and Wagener, 2008). One drawback of these *in silico* tools is their tendency to overpredict, thus generating more metabolites than actually occur, another is that often their underlying data originates from databases that focus on pharmaceuticals such that their coverage of chemistry may be more limited. (Boyce et al., 2022) and (Groff et al., 2024), have evaluated the performance and coverage of a number of these tools to provide some perspective of their utility to support read-across purposes. (Yordanova et al., 2021) provided a roadmap for how to use metabolic information from the QSAR Toolbox to substantiate read-across within categories, whereas (Gadaleta et al., 2020) provided a workflow for how predicted metabolism information could be used to identify and evaluate analogues for read-across. Key from these studies is the potential that alternative approaches can play in characterising metabolism, from the likely metabolites that can be formed, the sequence of transformations as well as the metabolic pathway itself. These aspects provide a means of codifying metabolism in a manner to permit its quantitative comparison. Some of these concepts were explored in (Yordanova et al., 2021; Boyce et al., 2022; Lester et al. 2023; and Patlewicz et al., 2024b) to showcase practical means by which metabolic similarity could be characterised to facilitate pairwise comparisons.

6.4.5. Considerations when applying a metabolic rationale

Some metabolic processes are ubiquitous and so is well understood that these can be presumed to occur without performing *in vivo* experiments in every instance. It may still be important to understand the kinetics of the metabolism even for well understood processes. A question remains over what data would be sufficient to demonstrate the existence and rate of the metabolic pathway to support the use of read-across. There are no objective thresholds to specify what constitutes ‘rapid’ metabolism, ‘predominant metabolite’ or whether the burden of evidence should differ depending on the endpoint itself or the absence/presence of effects, e.g. is there a different threshold for ‘rapid’ metabolism if the endpoint is for a developmental toxicity endpoint. That said, fixed and rigid criteria to establish threshold would not be desirable nor necessarily helpful.

The significance of toxicity of any residual parent chemical may also need to be considered. Exploring differences in toxicity and distribution (if any), where data on both parent and metabolite exist for the same endpoint can be helpful in this type of determination. Once a metabolic pathway has been demonstrated, if the toxicological profiles of the category members can be shown to be consistent, then the use of any read-across should be strengthened. If the approach was extended to include the consideration of other structurally related ‘metabolic pairs’ as part of a category approach, then this could also aid in reducing the uncertainty associated with the read-across proposed (Patlewicz et al., 2013b). In addition, when a simulation of metabolism is used in the analysis, adequacy of the simulated metabolism needs to be supported by experimental or theoretical data, to ensure a reliable assessment. Such an approach is proposed by Dermen et.al., 2022) where different aspects of similarities associated with metabolic degradation are considered in the estimation of reliability. The approach is particularly important for

confirming negative predictions, where it should be demonstrated that no metabolites associated with structural alerts or positive toxicity data are formed as a result of metabolism.

Practically, assessing the available ADME/TK data (the generation of which has been discussed above) in the light of the entire toxicological dataset, including information on the mode of action of the chemical, should provide a measure of the validity in the metabolic pathway approach. The US EPA High-Throughput Toxicokinetics (HTTK)[105] project has developed generic models and chemical-specific data for simulation and statistical analysis of chemical toxicokinetics as described by (Pearce et al., 2017).

The metabolic pathway approach is usually limited to systemic toxicity endpoints. Other endpoints of hazard identification studies that are dependent upon site of contact effects (e.g. eye, skin, respiratory tract irritation, irritation to gastric mucosa) cannot typically be addressed using the metabolic category logic. These sites of contact effects are often due to the physicochemical properties of the chemical in question and therefore may differ considerably between the parent compound and primary and secondary metabolites.

6.4.6. Developing a metabolic pathway category in practice

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

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

- Step 3 *"Develop category hypothesis to identify category members"*: Provide definitive information on the metabolism of the parent chemical to the primary and secondary metabolites. This information may include time course data for either blood or tissue for both the parent chemical and the primary and secondary metabolites. Alternatively, if appropriate, *in vitro* metabolism may generate sufficient information to support chemical grouping.
- Step 4 *"Gather data for each category member"*: The metabolism information should be examined to determine, if in fact, the primary and secondary metabolites are formed, if these achieve appreciable levels within the blood and/or tissues and determine basic toxicokinetic parameters for the parent material. If the metabolism of the parent chemical to the primary metabolite is rapid and is thought to occur within barrier tissues (e.g. lung, gut, placenta tissue), then it may be appropriate to use hazard identification studies from the primary metabolite to identify hazards associated with exposure to the parent chemical.
- Step 5 *"Construct a matrix of data availability"*: A quantitative analysis between exposures of the parent chemical and the primary and secondary metabolite is usually not necessary if the only objective is hazard identification. It is recognised that in certain cases quantitative differences can play an important role in hazard identification (e.g. in the metabolism of C6-C8 alkanes). For risk assessment purposes, a quantitative analysis may become necessary, e.g. additional ADME/TK analysis (including preparing a model) may be appropriate.
- Step 6 *"Evaluate the category members":* If there are appropriate hazard identification studies that have been conducted with the parent chemical or primary or secondary metabolites for similar toxicity endpoints, then these studies should be examined to see if these materials have similar toxicity. If data are not available for the metabolic series in question, then structurally related

[105] https://cran.r-project.org/web/packages/httk/index.html

metabolic pairs should be considered. If such information were not available, a study could be designed and conducted. In this case the parent compound could be tested. Any toxicokinetic and metabolic experiments that provide the basis for the metabolic category should have robust summaries prepared and be included in the dossier for the parent chemical, primary and secondary metabolites. A table should be included detailing the relative levels of the parent chemical, primary and secondary metabolites.

The metabolic pathway approach may not be applicable for environmental toxicity endpoints unless the metabolism of the parent compound to the primary or secondary metabolite can be demonstrated within the test species in question. Whereas it may be appropriate to extrapolate within mammals, it may not be appropriate to extrapolate directly from rodents to fish or between amphibia and fish or insects and other species due to the difference in the metabolic processes and enzymes present within those species. In addition, significant differences in metabolic capacity occurs often between different life stages and this should also be taken account of when relevant.

The same concept underlying the metabolic pathways can also be used for environmental degradation processes. For example, for a substance which hydrolyses very rapidly in aquatic test systems (half-life < 1 hour), the aquatic toxicity endpoints can be covered by the test results with the degradation product(s) (OECD, 2019c). A biotransformation/degradation pathway approach making use of biodegradation and metabolism studies can be useful to help in characterising bioaccumulation potential. If a parent substance was extensively degraded and/or metabolised, this type of information could be helpful to rationalise the likely bioaccumulation potential of a chemical. However, it should be considered whether the biotransformation pathway and rate observed is likely to be relevant for the species in which the bioaccumulation is considered, because it is well known that biotransformation is often highly dependent on taxonomic group and life stage. It should for example be carefully considered how much standard biodegradation studies such as ready or simulation biodegradability studies including analysis of the transformation products formed can really provide pertinent information to substantiate the read-across proposed if this is considered relative to the BCF in fish.

6.5. Substances with more than one constituent and mixtures

6.5.1. General considerations on substances with more than one constituent and mixtures

Most generally, individual substances[106] (as opposed to mixtures) are considered for grouping and read-across. These may, however, contain several constituents and/or impurities that need to be taken into consideration for grouping and read-across.

Multi-constituent substances are discrete substances where all constituents are well defined. On the other hand, in the case of UVCBs it may not be possible to fully characterise their chemical composition due to the large number of constituents, poorly defined/partially unknown composition and/or a high variability in the composition including variable concentrations. In many cases, chemical reaction products are UVCBs (see Section 6.6).

(OECD, 2018c) defines mixtures as *"co-existing set of two or more substances in which they do not react"*. Overall, mixtures can have a wide range of number of components, as well as being more or less well

[106] Globally Harmonised System of Classification and Labelling of Chemicals (GHS) defines a substance *as "including any additive necessary to preserve the stability of the product and any impurities deriving from the process used, but excluding any solvent which may be separated without affecting the stability of the substance or changing its composition"* (GHS, Tenth revised version, 2023; GHS/Rev.10e (https://unece.org/sites/default/files/2023-07/GHS%20Rev10e.pdf).

defined or variable, also concerning concentrations. Intentional mixtures are for example manufactured formulations. For mixtures of discrete substances, similar to well-defined multi-constituent substances, the nature and concentration of all components is known.

Categories can be developed for series of chemical reaction products (UVCBs) or multi-constituent substances that are related in some regular fashion.

However, the following aspects add complexity to category development and read-across for multi-constituent substances and UVCBs (see for example (ECHA, 2022b; ECHA, 2017c); see Section 6.6 on UVCB substances), and apply also to mixtures:

- Variations in constituents and their concentrations of the considered substances (or components of mixtures).
- There may be practical or technical limitations to fully characterising the compositions of UVCBs for the purposes of read-across e.g. to establish structural similarity.
- Possible influence of the constituents/components on each other's (eco)toxicity in combined exposure to multiple chemicals ("mixture toxicology").

Therefore, establishing similarity might be challenging. The composition and physicochemical properties of substances are useful considerations to take into account when dealing with multi-constituent substances. Two types of approaches can be considered (or a combination thereof):

- A constituent-based approach, considering properties of and using data available for individual constituents.
- A whole-substance approach, considering the properties of and using data obtained with the substance as a whole.

A number of categories assessed under the OECD Cooperative Chemicals Assessment Programme provide useful case studies on dealing with multi-constituent substances[107]. For the Ethylene Glycols category, data from PEG 200, a mixture of chain lengths, was used to support the human health assessment. For the linear Alkylbenzene Sulphonates category, aquatic toxicity data was available for both commercial products (C10-C14) and pure C13 and C14 homologues. The pure homologues showed higher toxicity than the commercial mixtures but data for the pure homologues was not used to drive the recommendation of the assessment since these were not commercially supplied (Caley et al., 2007). The Bicarbonate Special category focusing on ammonium bicarbonate, provided an interesting example of assessing a reaction mixture using data from pure components. The commercial material is a reaction mixture of sodium bicarbonate, sodium carbonate and ammonium bicarbonate. Aquatic toxicity data was available for the three components. Ammonium bicarbonate is the most toxic and the evaluation therefore focused on the quantity of ammonium ions released to water from dissolution of Bicarbonate Special and the impact of pH on the ammonium speciation and toxicity (Caley et al., 2007). Effectively, the ammonium ion was used as a marker for aquatic toxicity (see also Section 6.4).

6.5.2. Mixtures of discrete substances: Use of toxic equivalency factors or toxic units approach

Hazard assessment of chemical mixtures

The use of toxicity equivalency factors (TEFs) is a special case of the relative potency factor (RPF) method. It is reserved for the estimation of toxicity of mixtures of chemicals when evidence supports the assumption that all components induce a biological effect through an identical toxicity pathway. The TEF technique is applied under the strictest interpretation of dose additivity in which component-based dose response data

[107] Further information is available at https://hpvchemicals.oecd.org/ui/ChemGroup.aspx

are integrated to express the mixture's toxicity as a single value. Considering the requirement that the component of the mixture act via an identical toxicity pathway, this TEF approach is strictly applicable to mixtures that have been formally grouped based on detailed characterisation of mechanistic steps from molecular initiating event to terminal key event(s). Furthermore, toxicity data for the endpoint being assessed must be available for each component in the mixture.

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

Complex mixtures of polychlorinated biphenyls (PCBs) (Clemons et al., 1997), furans (Muir et al., 1992), dioxins (Safe, 1991; Van der Weiden, 1992) and aromatic hydrocarbons (Walker, 1991; Zabel, 1995; ATSDR, 2022) have been assessed using toxicity equivalency factors based on relative Ah receptor binding potency/affinity and joint toxicity models. Joint toxicity models for calculating TEFs generally use a strict dose addition model when a common toxicity pathway is demonstrated across mixture component chemicals. Although synergistic effects are conceivable, these are only observed when chemicals in a mixture have different mechanisms, which should not be the case within a chemical category rigorously formed by the same toxic MOA- considerations.

In the TEQs approach, the most toxicologically well characterised compound is used as the reference or index compound. This compound does not necessarily have to be present in the mixture being assessed, but the components of the mixture of interest must all act by the same MOA, have dose-response data for a common effect, and share structural similarity e.g. functional group(s) with the index compound. Under these conditions, a TEF is calculated for each mixture component by dividing the effect value (e.g. benchmark dose or effect dose) of the index compound by the effect value of the nth mixture component chemical (Box 6.2).

Box 6.2. Equation 2. Calculating TEF

$$TEF = \frac{Index\ compound\ effect\ value}{nth\ mixture\ component\ effect\ level\ value}$$

A TEF then represents the potency of a given mixture component relative to the potency of the index chemical for the selected health effect. To obtain an index chemical equivalent dose (ICED), the dose or concentration of each component in the mixture is then multiplied by its respective TEF. The ICED values for each component are then summed to give the overall toxic equivalency for the mixture (Box 6.3).

Box 6.3. Equation 3. Calculating TEQ

$$TEQ = \sum(concentration\ of\ nth\ mixture\ component\ x\ TEF\ of\ nth\ mixture\ component)$$
$$= \sum(ICEDs\ for\ all\ mixture\ components)$$

While the above equations represent the general approach to develop TEFs, more recently the concept of "Best-Estimate TEF" has been introduced that makes use of machine-learning based quality weighting of REPs and Bayesian dose-response modelling and meta-analysis (DeVito et al., 2024).

For example, in the case of dioxin and furan mixtures, toxicity of components relative to 2,3,7,8-tetrachloro-p-dioxin (2,3,7,8-TCDD) was derived from the updated 2004 relative effect potencies (REP) database, using a consensus-based weighting scheme, a Bayesian dose response modelling and meta-analysis to derive "Best-Estimate" TEFs (DeVito et al., 2024). The following table (Table 15) lists TEFs for four dioxins using *in vivo* and *in vitro* studies:

Table 15. Toxic Equivalency Factors (TEFs) for `four dioxins

Dioxin	TEF
2,3,7,8-tetraCDD	1 (reference compound)
1,2,3,7,8-pentaCDD	0.4
1,2,3,7,8,9-hexaCDD	0.05
1,2,3,6,7,8-hexaCDD	0.07

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

- Mixture A contains 20% 2,3,7,8-TCDD, 50% 1,2,3,7,8-pentaCDD, 10% 1,2,3,7,8,9-hexaCDD and 20% 1,2,3,6,7,8-hexaCDD. To calculate the joint toxicity of the mixture for this specific effect, the respective component fractionation data and TEFs are leveraged in Equation 2:

  (0.2 x 1) + (0.5 x 0.4) + (0.1 x 0.05) + (0.2 x 0.07) = 0.419

- In order to translate this TEQ into an effect value for Mixture A, the effect-specific value of 2,3,7,8-TCDD (e.g. ED_{10}) from the selected dose-response/study is divided by 0.419; the interpretation is that the joint toxicity of the four components together, acting via an identical MOA, results in a greater effect dose- response.

TEF/TEQ approach is used for example in the regulatory context of food safety for dioxins and polychlorinated biphenyls[108].

The TEF/TEQ approach has been adapted for neurotoxicity with a relaxation of some of the common MOA constraints in order to facilitate hazard assessment of PCBs. (Pradeep et al., 2019) extended work by

[108] see Commission Regulation (EU) 2017/771 of 3 May 2017 amending Regulation (EC) No 152/2009 as regards the methods for the determination of the levels of dioxins and polychlorinated biphenyls

(Simon et al., 2007) to explore the derivation of neurotoxic equivalent factors (NEFs) using congener potency data from 7 different *in vitro* studies. The assays measured protein kinase C (PKC) translocation, microsomal and mitochondrial calcium sequestration, dopamine content reduction, ryanodine receptor type 1 activity, inhibition of vesicular transport-mediated uptake of dopamine and glutamate, inhibition of transport-mediated uptake of dopamine and PCB interference with [3H] WIN 35,428 binding at DAT. Since not all congeners were tested in all of the assays, a neurotoxic relative potency value (REP) was first derived which was then used to derive an average NEF per tested congener. (Q)SAR models were then derived using structural features characterising the PCBs in particular capturing the substitution pattern of the chlorines so that predictions of NEFs for the untested congeners could be made.

In summary, toxic equivalency can be used for complex mixtures for which there is a presumed common MOA such that the doses of the individual components (discrete substances) taking their potency info into account, is additive. Toxicity data should ideally be available for each individual component of the mixture. In absence of empirical data, predictions of toxicity potency from (Q)SARs may be feasible per the example highlighted in (Pradeep et al., 2019). It should be recognised that differences in test protocol for each data point can have a marked effect on the derived component TEFs (and so mixture TEQ), therefore if this approach is followed then it is necessary to present all available data and justify the use of the approach.

Generalising the TEF to assess the similarity of mixtures

The development and application of TEFs, NEFs, and RPFs can strengthen conclusions drawn from emerging methods for assessing "sufficient similarity" among chemical components in mixtures. When toxicity data are available for one or more "whole" chemical mixtures within a specific class, statistical approaches can be used to evaluate the extent of toxicological similarity between these mixtures and whether they are "sufficiently similar" for the purpose of risk assessment application (e.g. whether one mixture's dose-response data could be used as a surrogate in the risk assessment of another target mixture).

The development of methods for evaluating sufficient similarity is an active area of research, and various approaches have been applied to diverse chemical groups, including pyrethroids (Marshall et al., 2013), botanical supplements (Catlin et al., 2018; Ryan et al., 2019), petroleum substances (Murray et al., 2013), and wildfire smoke (Koval et al., 2022). There are also ongoing efforts to develop toolkits and training materials to support mixtures similarity analyses (Roell et al., 2023) and to consider how recent advances in non-targeted chemical analyses, new approach methodologies for profiling bioactivities, and other computational approaches and similarity indices can be used to inform similarity evaluations (Rager and Rider 2023; Roelle et al., 2023). For example, US EPA is developing approaches and tools to evaluate sufficient similarity of PCB mixtures for use by risk assessors to determine appropriate surrogates when mixture-specific data are not available[109]. EFSA has developed an approach for grouping pesticides into cumulative assessment groups (CAGs)[110] based on their similar toxicological properties in a specific organ or system to inform cumulative risk assessment.

6.6. Unknown or Variable composition, Complex reaction products or Biological material (UVCB substances, or UVCBs)

The main issue for grouping of UVCBs (e.g. petroleum products, resins, essential oils, and other natural complex substances) is that the composition of UVCBs is poorly defined and constituents are highly

[109] https://cfpub.epa.gov/ncea/iris_drafts/recordisplay.cfm?deid=354040

[110] https://www.efsa.europa.eu/en/press/news/efsa-presents-cumulative-assessment-group-methodolog

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

Generally, grouping and read-across for UVCBs can be constituent-based and/or consider properties of the UVCB substance as a whole[111].

UVCB substances, which by definition may contain multiple and variable constituents, are complex substances that may be grouped together if they are expected to contain similar constituents and thus exert similar effects. The range of different types of UVCB is very wide and the specific properties may be diverse, so the justification for using a category approach should address the compositional variability of the UVCB substance and the impact on the properties of interest. There are many different types of UVCBs, although generally these all have the following characteristics in common:

- These contain numerous constituents and cannot readily be represented by a simple chemical structure or defined by a specific molecular formula.
- These are not intentional mixtures of chemicals.
- Many are of natural origin (e.g. crude oil, coal, plant extracts).
- Many cannot be completely separated into their constituent chemical species.
- Some are defined by their starting materials and their manufacturing process.
- These are often produced to meet a performance specification related to their physicochemical properties.
- The concept of “impurities” does not apply to UVCBs.

While CAS Registry Numbers are important for identifying substances, in the case of UVCBs, the specificity of the substance, name and manufacturing process description, starting material are important for UVCB identity and can vary considerably between different CAS numbers. For UVCBs, CAS number definitions are not always available. If available, these definitions can be narrow or broad e.g. CAS numbers for:

- Petroleum substances are based on a number of considerations including chemical characteristics such as hydrocarbon type, carbon number range, content variability of (poly)aromatic, aliphatic naphthenic, aliphatics and S and N containing hetero-cyclic constituents. In addition, production and processing characteristics such as distillation range as well as the last processing step and physicochemical properties such as viscosity may be included as this information provides essential insight in the characteristics of constituents potentially present.
- Coal derived complex substances are typically based on the applied production process and may include information on the distillation range and the chemical composition.
- Natural complex substances (NCS), like essential oils, are assigned based on their genus and species, in some cases part of plant, extraction method and other processing descriptors (Api et al., 2022).
- Complex inorganic substances usually qualify as UVCBs (Rasmussen et al., 1999; US EPA 2005 a and b), due to their complexity and variability in composition and physical form, containing varying amounts of metals, metal compounds and/or minerals. These may occur naturally (e.g. mineral ores and concentrates) or be manufactured during the various refining streams of the metal and mineral industry (e.g. metal intermediates). Such substances cannot be sufficiently identified by parameters like the International Union of Pure and Applied Chemistry (IUPAC) nomenclature and/or other name or identifiers or by its molecular, structural information, or

[111] In the remit of EU REACH, considerations on multi-constituent substances and UVCBs are detailed in (ECHA, 2017c), advice on using read-across for UVCB substances is available at: https://echa.europa.eu/documents/10162/11395738/advice_uvcb_read-across_en.pdf/ac1f64a6-9ee5-441e-cf1c-92914b843b4e?t=1651665130365 (ECHA, 2022b).

chemical composition. Their description typically includes the origin of the substance, the production process (if applicable) and the composition (see further in Section 6.7).

Due to these numerous considerations, and for historical and geographical reasons, similar substances can have different CAS numbers. UVCBs with the same CAS number (as defined by their starting material and manufacturing process) can have significantly different properties and therefore can be part of different categories depending on their constituent composition or specification.

These complexities have sometimes led to the use of physical properties and chemical descriptors (e.g. chain length, chemical class, size of aromatic ring systems) or substance composition as a way to define categories of UVCBs. The OECD developed guidance on how oleochemical substances can be characterised in a way that their composition is accurately and consistently described for hazard assessment purposes (OECD, 2017e). In the case of NCS, this categorisation may also occur around the major chemical constituent(s) present and might include marker chemicals for toxicity when it is clear that the behaviour of the UVCB substances is driven by those marker chemicals.

6.6.1. General guidance on developing categories for organic UVCBs

UVCBs pose challenges for hazard evaluation and for judging the similarity between these complex substances and adequacy of read-across for data gap filling. The justification of a category or read-across argument will need to include some demonstration of the compositional or structural similarity of the category members or analogues for read-across in order to support the supposition that the category members/analogues have similar properties; indeed, a quantitative and qualitative comparison of the actual composition between UVCB substances is mandatory in some schemes (e.g. EU REACH[112]).

The critical issue when considering UVCB substances is composition, including the variability of the composition e.g. from batch to batch[113]. In order to determine the viability of using read-across or to define a category, one needs to understand the constituents of these substances in sufficient detail, which is a well-established challenge (Salvito et al., 2020; Salvito et al., 2022). It is also desirable to determine which of these constituents are likely to drive potential effects (e.g. benzene, 1,3-butadiene, polycyclic aromatic hydrocarbons (PAHs)). Criteria for category definition may include concentration range and typical concentration of constituents of the UVCB including what are the generic constituents; what are the specific constituents; what are acceptable constituent concentration ranges; as well as specifying exclusion criteria for category membership.

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

For organic UVCBs, it is important to clearly characterise the identity of the constituents and the composition of the complex substance to the extent it is relevant for hazard characterization. A meaningful indication of variability should be provided. In particular, it is necessary to identify which of the following attributes are key and must be specified:

- Starting materials(s) source and composition, e.g. for biological NCS identification of the genus/species origin should be considered.
- Information on the production process, e.g. distillation temperature range, catalytic processes.
- Known or generic composition description:

[112] COMMISSION REGULATION (EU) 2021/979 of 17 June 2021 amending Annexes VII to XI to Regulation (EC) No 1907/2006 of the European Parliament and of the Council concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH)

[113] For example, in the remit of REACH, measurements should be made in at least five independent samples from different production batches, see ECHA Advice on using read-across for UVCB substances (ECHA 2022b).

    - Range of chain lengths or predominant carbon number range or size of condensed ring systems
    - Marker constituents (such as aromatic content), if appropriate they should be clearly identified and if possible quantified for all category members.
    - Elemental composition, e.g. halogens, metals
- Physicochemical characteristics:
    - Standard index, e.g. colour index number.
    - K_{ow} ranges.
    - Chromatographic and other physical "fingerprints".
- Relevant product specification:
    - Cut off ranges, e.g. for boiling point.
    - Reference to standards.

It is possible to read-across between UVCB substances if the similarity for the specified endpoints can be adequately demonstrated. Alternatively, if all or most of the constituents of a UVCB are similar or fall within an expected range for an endpoint, then a single representative constituent or set of constituents can be read across to the whole UVCB. In this case, it is necessary to identify representative constituents of the UVCBs to cover the carbon range and chemical families of the UVCB, making sure to address any outlying properties (e.g., specific toxicity of hexane compared to other aliphatic hydrocarbons, higher water solubility of aromatic hydrocarbons compared to aliphatic hydrocarbons). This approach is explained in more detail below. Specific requirements might need to be taken into consideration depending on the applicable legislation.[114]

Compositional information for UVCB substances to support category/read-across approaches

The chemical space of a UVCB, which may encompass a large number of individual constituents, often prevents their full enumeration. The ability to generate representative structures and predict their hazardous properties can be used in hazard and risk assessment of UVCB substances by pointing out the regions of chemical space of greatest concern. Targeted analytical characterisation of constituents of representative structures can be determined in some cases, and the knowledge of constituent concentration(s) can inform the hazard identification of UVCB substances as well as the category justification and read-across for UVCB substances. If there are known constituents which would drive the endpoints of interest, it is possible to define a category based on the presence of those constituents. Furthermore, the presence of classifiable constituents (e.g. benzene) can inform on the classification of the UVCB.

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

UVCBs encompass many types of substances, from those containing a handful of constituents to thousands of constituents. The following considerations on compositional information are given based on examples for petroleum substances and other complex hydrocarbon UVCBs, for which composition characterisation is a particular challenge. For less complex UVCB substances with hundreds of constituents, standard industry methods, including chromatographic techniques (gas, and liquid

[114] For the remit of REACH see ECHA, 2017c, 2022b.

chromatography) and mass spectrometry, may provide adequate compositional information (Concawe, 2019). For middle-distillate substances with thousands, to hundreds of thousands, of constituents, two-dimensional gas chromatography (GCxGC), may provide significant compositional insight into the substance. However, a high amount of effort and expertise is required to develop methodology and interpret the results (Concawe, 2019; 2022). The provision of comprehensive qualitative and quantitative information on constituents present in the heavier petroleum UVCBs remains a challenge.

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

The alkyl chain may differ in the length, the number of (conjugated) double and triple bonds, the degree of branching, the presence of aromatic and non-aromatic rings, and the position of the functional group(s). The position of unsaturated bonds may be limited to certain parts of the chain, as in alpha-olefins. Branching may also be limited to certain positions of the chain with respect to the unsaturated bond, such as the vinyl, allyl, or at carbon atoms further from the unsaturated bond. The alkyl branches may have odd, even or arbitrary number of carbons, depending on the source of the starting material, i.e. of natural or synthetic origin, and the process. The alkyl chain may be defined with a generic description, such as tallow (animal fat, contains primarily glycerides of C16-18 fatty acids).

It is also possible to simplify the composition of UVCBs based on representative structures (see also below). The representative structures approach is mainly applicable to organic UVCBs, such as hydrocarbon solvents, as well as oligomers. One approach developed for petroleum substances but expandable to other hydrocarbons is the Hydrocarbon Block method (CONCAWE, 1996) in which the substance is divided into fractions or “blocks” defined by carbon number and chemical class. It is possible to use an analytical method like two-dimensional gas chromatography flame ionization detector (GCxGC-FID) to assess the concentration of each block for a given substance. Inorganic and organometallic UVCB substances are however more difficult to handle in this way and are discussed further in Section 6.7.

Category justification based on representative structures

The key step is to define the category and identify category members. While initially this may seem repetitive, in fact the steps are different for UVCBs. This is best explained by considering the “define analogue(s)” step, which for UVCBs means identifying single constituents that represent the range of properties and the matrix being built up by the UVCBs. The repetitive changes in the constituents may include:

- The length and branching of the hydrocarbon chain.
- Presence and number of aromatic rings.
- Presence, number, and position of different functional groups.
- Presence and position of heteroatoms, or different forms of isomerism.

The systematic generation of representative structures through combinatorial algorithms is especially useful when screening for outlying behaviour of the constituents. Presence or absence of such constituents should be specifically checked and should be included in the description of the substance. The presence might also be included in the name of the substance, together with a quantitative indicator (e.g. > 2% aromatics). The screening in certain cases, like presence of alerting functional group, could be done on

generic chemical description (e.g. by SMARTS[115] or Markush-type structure[116]). However, in most of the cases, full enumeration would be required. It is important to know the type of variations in structure that the complex substance could cover, and to avoid ignorance of potentially dangerous (classes of) constituents due to limitations in the multiplication algorithm. Some tools offer random combinations with high coverage of the theoretical chemical space, in case the number of the possible variants is too large. Therefore, tools that allow such enumeration need to be applied, followed by computational screening.

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

In situations where detailed constituent analysis is not possible due to the availability of currently available analytical technology, combining analytical information with detailed bioactivity similarities and/or biological MOAs can support category grouping (Lai et al., 2022; House et al., 2022; Grimm et al., 2016).

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

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

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

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

The computational demand in the multiplication and screening of organic UVCB constituents and the practical issues (Lai et al., 2022) in handling these processes had led to a proposal for the integration of the developed methodologies in a single software tool.

Data gap filling – testing

Where it is necessary to identify representative UVCBs for testing purposes, this should be done bearing in mind the key constituents of the category definition and the ranges thus defined.

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

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

Good practices in developing categories for complex substances (UVCBs)

In forming chemical categories made of UVCBs, the following good practice should be observed to enable hazard assessment and provide a sound grouping and data gap filling approach. It should be noted that specific requirements might apply depending on the specific regulatory decision context.

As for all substances, the clear identification and characterisation of the source and target substances is a pre-requisite for grouping (see above for a list of key attributes to be specified). This includes consideration of the following for UVCBs:

- Distinguish the individual constituents as far as possible and adequate, providing compositions and variability ranges as narrow as possible.
- Name the substance in a clear and consistent way, which reflects both the constituents and the composition, in accordance with relevant and existing competent authority requirements.
- Define the substance in an unambiguous way, with regard to inclusion or exclusion of (groups of) constituents, or by provision of a boundary composition.
- Support the identity and composition of the substances with analytical data to enable comparison with the category definition.

Furthermore, as for other types of substances, it is important to:

- Base the grouping on a scientifically credible and verifiable hypothesis;
- Subgroup the category, if there is a scientific justification.
- Justify the mechanistic rationale for the category if possible (per endpoint, if necessary).
- Explain, how the read-across is being made for the various endpoints to fill data gaps.
- Before attempting hazard assessment, make sure that there is sufficient data to allow trend analysis where applicable and there is a good coverage of chemistry and properties within the (sub)category, otherwise consider testing.

Other aspects to consider:

- Identify representative constituents.
- Identify constituents with known hazard.
- Consider possible interactions of constituents in terms of hazard.
- Analyse the constituents and substances for outlying behaviour.
- Identify reasonable worst-case scenarios for chemical hazard by endpoint, for constituents in the substance, and for substances in (sub)category, and ensure these are reasonably addressed.
- Consider variability in concentrations.

To approach grouping UVCBs in a step-wise fashion, the following steps can be considered, acknowledging that the grouping approach depends on the types of substances and the regulatory requirements:

1. Provide all analytical, physicochemical data (available/that can be generated as it is technically possible) and broad definition of manufacturing information on each complex or UVCB substance.
2. On the basis of the information in the above point, generate representative structures for all substances in the category, if possible (e.g. the Hydrocarbon Block method (CONCAWE 1996), UVCB methodology for structural description (Dimitrov et al., 2015)).
3. Merge the representative structures for all substances in the category in one pool that covers the chemistry spanned by the whole set of UVCB substances in the category.
4. Collect all available information for the representative structures, including data from experimental databases or predictive methods, for the endpoint(s) for which there is a data gap.

5. Group the representative structures in groups that have similar hazard profile (or fate property depending on the endpoint that has the data gap).
6. Build an analytical matrix that shows the mass fraction of each UVCB substance for each identified group of representative structures (e.g. Hydrocarbon Block method (CONCAWE 1996) to petroleum UVCB identification) or build an analytical matrix where substance identity is known, containing each constituent; if concentration is unknown and there is significant variation in the profile of the different groups of representative structures, this should be evidence that the category cannot hold without further analytical characterisation. One approach to support the category approach where analytical characterisation is insufficient, grouping may be supported using biological endpoints or bioactivity similarities.
7. Model, using as independent variables, the analytical composition of the UVCB substances expressed as the mass fraction of the UVCB. Or each group of representative structures that has a significant hazard profile, if possible.

Points 4-7 will need to be repeated for each endpoint. The category may stand for some endpoints but not for others (e.g. a category can stand for systemic but not topical effects, or vice versa). The computational approach depends also on the computational possibility to predict endpoints.

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

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

6.6.2. Hydrocarbon solvents

In order to develop a category-based read-across strategy for hydrocarbon solvents, understanding their composition and manufacturing history is crucial to avoid confusing them with other less refined petroleum substances with comparable carbon numbers such as mineral oil.

UVCB hydrocarbon solvent substances are commonly derived from petroleum or alternative source of feedstocks and typically contain the following types of hydrocarbons:

- n- paraffins – N – saturated linear hydrocarbons (n-alkanes)
- iso-paraffins – I – saturated branched hydrocarbons (branched alkanes)
- cycloalkanes – C – saturated cyclic hydrocarbons (cycloparaffins) including alkyl sidechains
- aromatic – A – aromatic hydrocarbons, including alkyl side chains

The constituent distribution and relative presence in hydrocarbon solvents is referred as “NICA” composition. Hence the difference between the types of hydrocarbon solvents is mainly due to their relative difference in hydrocarbon classes and their carbon number ranges. The carbon number distribution is usually narrow (~ 3 to 5 carbon numbers), defined by the distillation range, and typically and nominally between C5 and C20, with a constrained NICA composition.

By convention, the carbon number range is based on the alkane and aromatic constituents' distribution. Aromatics have substantially higher boiling points than alkanes with the same carbon numbers. For aromatics the boiling point goes up in association with the increasing carbon number and with increasing ring number (2-ring C14 aromatic: 292°C; 3-ring C14 aromatic: 343°C). Therefore, hydrocarbon solvents with carbon numbers up to C20 contain generally aromatics with lower carbon numbers, mostly alkylated one and/or two ring species.

The major processes for transforming all types of feed stocks into hydrocarbon solvent substances include distillation of the feedstock, hydrodesulphurisation, mild or heavy hydrotreatment and finally distillation and stripping of light constituents. These processes are designed to deliver solvents that consists almost exclusively of hydrogen and carbon (e.g. removal of sulphur), thus "hydrocarbon solvents". More recently, an increasing number of synthetic hydrocarbon solvents are manufactured using natural gas, coal, or lower molecular weight hydrocarbons via some type of oligomerisation process. The manufacturing processes are designed to eliminate olefins and heteroatoms and keep aromatics such as BTEX (Benzene, Toluene, Ethylbenzene and Xylene) at low levels.

Hydrocarbon solvent groupings or categories rely on composition; typical NICA chemistry and carbon-number ranges, because the toxicological profiles of hydrocarbon solvents indicate these physicochemical characteristics drive their human safety profile and environmental effects. Most products manufactured from petroleum feedstocks have used process descriptors as the basis for CAS numbers. In fact, the use of CAS numbers for current hydrocarbon solvents has introduced confusion because they were developed (some 50 years ago) in a very generic sense for petroleum-derived substances and now tend to be overly broad and imprecise: CAS numbers are based on source and process information and do not accurately reflect the specific narrow composition and characteristics of hydrocarbon solvents which often leads to confusion between petroleum fuel-feedstocks and hydrocarbon solvents. Therefore, more precise composition-based descriptors have been developed enabling better grouping, risk assessments and regulatory compliance for hydrocarbon solvents. In the EU for REACH registration purposes, EC number are used, and a naming convention has been developed based on their NICA composition. This naming convention facilitates the qualitative description.

The following qualitative descriptors of the constituents present in hydrocarbon solvents should be provided to facilitate grouping by composition:

- NICA composition as described in "ii)" below
- Number of carbon atoms in the constituents (carbon numbers)
- Constituents with specific hazard profiles, particularly those with hazard classifications (e.g. naphthalene, n-hexane)

Furthermore, grouping of similar constituents is aided by applying a systematic naming convention to hydrocarbon solvents based on the following constituent descriptors:

i) Chemical character descriptor
ii) Carbon number descriptor (carbon range)
iii) Hydrocarbon structure descriptor
iv) Specific hazard constituent descriptor

- Chemical character descriptor identifies the chemical character of the substance as "Hydrocarbons", namely molecules constituted entirely of carbon and hydrogen atoms.
- Carbon number descriptor describes the number of carbon atoms in the carbon chain length(s), including the carbons in cycloalkanes: - In general, the carbon number descriptor refers to the overall major carbon number range of the alkane constituents irrespective of the hydrocarbon structure, e.g. "C12–C14" corresponds to "C12, C13, C14" including both even and odd numbered alkyl-chains. The carbon number range reflected in the name is

defined by sum of NICA constituents that represent at least 80% (w/w) of the total composition and individual constituents at levels above 10% (w/w) are also mentioned. Individual constituents at level below 10% (w/w) are not included in the naming of the substance unless they trigger a health or environmental classification (see point d., specific hazard constituent descriptors).

- If the carbon number descriptor consists of only even or odd numbered alkyl chains e.g. C12-C14 (even numbered), this should be indicated (see next point).

i. Hydrocarbon structure descriptor refers to the type of hydrocarbon structures – NICA present and are included after the chemical character (hydrocarbons) and carbon number descriptor (C12-C14) separated by a comma, e.g. "Hydrocarbons, C12–C14, n-alkanes, iso-alkanes, cyclics < 2% aromatics".

- The level of aromatics is indicated individually or in combination with alkanes constituents. When present with alkanes it is indicated as a maximum or in ranges for example: as <2% or 2-25% aromatics. So that a C12-C14 solvent with > 2% aromatics is named "Hydrocarbons, C12–C14, n-alkanes, iso-alkanes, cyclics, aromatics 2-25% ." or individually as "Hydrocarbons, C9 aromatics" for a solvent that is essentially only composed of aromatics of a given carbon number.

ii. Specific hazard constituent descriptor refers to constituents that cause systemic effects that differ from the common toxicological profile and are present at significant concentrations that may result in regulatory classification will be indicated (e.g. n-hexane and naphthalene). For example, "Hydrocarbons, C10, aromatics, <1% naphthalene" can be distinguished from a similar substance with higher naphthalene levels (> 1%) that is classified based on its naphthalene content; "Hydrocarbons, C10, aromatics, >1% naphthalene". In the case of n-hexane a 5% threshold is used for neurotoxic effects "Hydrocarbons, C6-C7, isoalkanes, cyclics, <5% n-hexane".

The naming convention helps also in distinguishing hydrocarbon solvents with complex compositions (UVCB solvents) from mono-constituent substances but may share a common toxicological profile to be grouped together for hazard and risk assessment purposes. The following qualitative criteria are used:

- 80% (w/w) rule. Applicable to the sum of NICA analysis or if one constituent is present at this minimum concentration to be considered a "mono-constituent" so that the hydrocarbon solvent can be named using a trivial name. For example, "n-hexane" and "iso-hexane" would indicate these are two distinct C6 solvents composed of > 80% n-hexane or iso-hexane.
- < 80%. When the sum of NICA constituents does not achieve 80% (w/w), all carbon numbers regardless of their concentrations should be considered for their characterization. Those present < 10% (w/w) may be omitted from the carbon number descriptor. If the carbon number or hydrocarbon type is unknown > 20% (w/w), it should be considered whether the UVCB substance can be regarded an hydrocarbon solvent.

These naming principles are explained in more detail in the "OECD Guidance for Characterising Hydrocarbon Solvents for Assessment Purposes" (OECD, 2015).

From a toxicological hazard assessment perspective, there is an extensive existing dataset which justifies an overarching read-across strategy (McKee et al., 2015) considering both carbon chain length and hydrocarbon type. This type of category approach uses the existing database to predict the toxicological properties of all hydrocarbon solvents based on the premise that all the substances in these categories have qualitatively similar properties; that ADME considerations including known metabolic processes and metabolites are understood, consistent and predictable; that classified constituents such as n-hexane and naphthalene which are present in some specific hydrocarbon solvents are quantified.

6.6.3. Coal derived complex substances

The principle described in Section 6.6.2 for petroleum derived complex substances also applies to coal derived complex substances. The longer geological history of coal compared to crude oil explains the higher degree of cross-linking of coal derived constituents. This results in a predominance of aromatic ring systems in coal derived complex substances. Longer alkyl chains do not appear. Processing of a coal derived feedstock separates according to volatility (size of condensed ring systems) and/or the extractability of acidic/ alkaline constituents. Formation of categories makes use of the applied processing techniques and of a similar spectrum of intrinsic properties for substances having a similar matrix of physicochemical properties.

6.6.4. Natural complex substances (NCS)

NCS can originate from plant, animal, or microorganisms. Some inorganic UVCBs are also natural substances e.g. natural clay minerals. For example, NCS include botanically-derived substances obtained by subjecting specific parts of the plant to a physical treatment such as extraction, distillation, expression, fractionation, purification, concentration or to fermentation. Their compositions vary depending on the genus, species, the growing conditions, and maturity of the crop used as a source, and the process used for its treatment.

NCS constitute a very specific subgroup of UVCBs and include primarily essential oils and extracts obtained by various separation techniques.

Inclusion in a chemical group is possible based on the constituents of the NCS where the major constituents can be clearly identified as the same as known chemical substances. A full workflow outlining the approach that this is performed for fragrance substances is described in (Api et al., 2022). Herein, NCS are first identified by plant taxonomy: Family, Genus, and Species. Further identifiers, including plant part and extraction processing methodologies, are used to clarify substance identification. The chemical similarity of complex mixtures then considers two distinct aspects; whether the constituent structural identity is the same across NCS ("structural identity" meaning the number of constituents common versus uncommon), and whether compositional percentages of the constituents are the same.

Ores and concentrates are also naturally occurring substances of variable composition and are discussed under Section 6.6.5.

6.6.5. Developing categories for complex inorganic UVCB substances

Complex inorganic UVCB substances contain varying amounts of metals, metal compounds and/or minerals in different chemical forms ('speciation'). These substances may occur naturally (e.g. mineral ores) or be generated and used during the primary and secondary production (recycling) of metal substances, enamels, ceramics, glass and inorganic pigments, and the refining streams of the mineral industry. Their (eco)toxicity is related to the released metal/mineral ions and their presence at target sites (Goyer, 1996).

Ores and concentrates

Ores are composed of metal-bearing minerals of sufficient quantity and quality to be mined for profit. Once mined, ores are physically processed to remove minerals of no economic interest ('gangue minerals'), allowing value minerals to accumulate within a concentrate. Ores and concentrates can be complex and are usually less well-defined than most substances: the natural variability in their mineral composition (even within the same ore body) and the various mitigating effects of their mineralogy and physical forms render substance identification for hazard identification purposes challenging. The full hazard profiles of ores and concentrates as 'a whole' are usually not available. Grouping and read-across is applied to

characterise the hazard and derive a classification, by using (eco)toxicological data for one or more “source substance(s)” (i.e. a representative ore or concentrate) with similar physical and chemical properties and a known (eco)toxicological profile.

Grouping of ores and concentrates can be done by combining the available information on the composition, the origin (location and processing) and the expected (eco)toxicological properties associated with the released metal/mineral ion. This release is dependent on a number of factors e.g. the composition, physical form, particle size, presence of a mineral matrix and the solubility of the minerals (carbonates, oxides, sulphides, etc.).

The mineralogical composition (provided by e.g. optic microscopy, X-ray diffraction or more recently by advanced X-ray crystallography, SEM-EDX, Mineral Liberation Analysis) is the key information to predict any release of an element from geological material that may cause toxicity but is often difficult to obtain fully beyond the major constituents. However, for these substances, chemical speciation data is often a cost-effective proxy and therefore more readily available.

In a first step, grouping is done by organising the materials into groups based on the main metal contained in the material (e.g. zinc concentrates). These groups can further be split in more specific groups, for example by considering, within the group of zinc concentrates, the materials originating from sulphidic zinc ores. Additional insights may come from experienced geologists, geochemists, and mineralogists familiar with both the ore body and concentration process. Transformation/dissolution tests (OECD 29, 2002), measuring the metal ion release to the environmental media, can provide further insights in the ‘behaviour’ of the substance (see also Section 6.7).

Ultimately, the grouping should result in a “functional group” of materials with expected similar physicochemical and (eco)toxicological properties using a weight of evidence approach. The justification of the grouping must be provided based on the physicochemical and toxicological properties used to determine the groupings. Within such a group, a generic assessment can be made on a carefully chosen “representative material” of the group, based on the available data for the substances in each group, instead of analysing every single specific material of the group. The hazard identification (and classification), for each human health and environmental endpoint of the ores/concentrates in the group will be based upon data available for the representative ore or concentrate when data for that ore or concentrate is not available. Alternatively, the hazard identification of each ore and concentrate within the group may be based on a read-across approach considering the concentration, properties, and hazard profile of the individual constituents. This approach is further described and implemented in the MeClas tool (www.meclas.eu; Verdonck et al., 2017d).

Inorganic UVCBs produced during production and recycling

Inorganic UVCBs are also generated and used during the non-ferrous metal production. They can be collected as by-products or removed down the process steps as enriched metal intermediates (i.e. rich in a specific metal but not yet a pure metal substance). These inorganic UVCBs are typically composed of several constituents, whose concentration will vary depending on the material used as input for the process as well as on the parameters of the process production. Examples of such materials include slimes and sludges, slags, matte, flue dust etc.

The variability of the inorganic UVCBs in composition, physical form/particle size, speciation, dissolution kinetics and the presence of a matrix determines the amount of metal that is released and may exert toxic effects. The high variability complicates fulfilling the normal (eco)toxicity assessment requirements for substances. Subjecting all inorganic UVCB substances to (eco)toxicity testing is hardly feasible. In addition, the test of a single UVCB of interest does not lead to results representative for the UVCB variability, as worst-case materials can vary depending on the endpoint under testing. Even these worst-case materials would not be applicable to all UVCBs but should be grouped by those of similar

physicochemical and toxicological properties before determining the worst case for that group for each endpoint.

A specific approach for such UVCBs has been discussed in a Metals and Inorganics Sectorial Approach (MISA)[117] workshop for the assessment of the inorganic UVCBs to appropriately account for their intrinsic variability and uncertainty. This has been applied in the EU and beyond. It starts with:

a) An assessment of 'substance sameness' based on a description of the (elemental, mineral) composition, the type of process and the type of input material in the process.
b) The identification of a boundary composition (known or reasonable worst-case).
c) The calculation of a conservative classification by assigning constituents' speciation and a worst-case composition (maximum of the typical concentrations for each constituent) that are applied to all the UVCBs that are part of each sameness group. The further risk assessment will be based on the constituents of the UVCBs, taking reasonable worst-case assumptions for hazard and exposure

To support and harmonise the hazard assessments/classifications of inorganic UVCBs, and hence also streamlining grouping and read-across, the industry developed a freely accessible, web-based tool (MeClas, www.meclas.eu) and continues to enhance it with new data and functionality. It addresses the specific challenges associated with the human health and environmental hazard assessment of such complex inorganic materials, intends to take into account the UN GHS, EU CLP and US OSHA requirements and rules. The tool includes links to (eco)toxicity reference databases and works by tiers considering the amount of available information (Verdonck et al., 2017).

6.7. Metals and inorganic compounds

6.7.1. Introduction

The concept of grouping has traditionally been widely used for hazard- and risk assessment of metal substances. For example, grouping and subsequent read-across has extensively been applied by industry and regulatory agencies as an alternative to animal testing to provide hazard information, including for classification and labelling purposes (e.g. harmonised classifications of > 100 inorganic nickel substances under the EU CLP).

The grouping approach has also been applied to estimate the potency of the effects by read-across with, for example, NOAELs and NOECs being read-across from data obtained from water-soluble metal compounds ('source substances') to other water-soluble compounds of the same metal (by same route of exposure and for same endpoint; 'target substances'). Examples include EU risk assessments on nickel substances (Tsakovska and Worth, 2007; ECB, 2008a) and zinc substances, which were endorsed by the OECD (ECB, 2008b; OECD, 2005b) or testing proposals under EU REACH for cobalt salts.

The Substances Grouping Initiative under Canada's Chemicals Management Plan (CMP) applied grouping approaches to carry out risk assessments (e.g. Cobalt-containing Substance Grouping, Selenium-containing Substance Grouping[118]). ECHA has proposed grouping of several metal substances in its assessment of regulatory needs (ARN), based on the intrinsic properties of the metal cation released from

[117] Metals and Inorganics Sectorial Approach (MISA), a cooperative programme set up by ECHA and Eurometaux, the European non-ferrous metals association. Workshop report: https://www.reach-metals.eu/uploads/pdf/MISA%203rd%20workshop%20November%202019/MISA%203%20Executive%20Summary.pdf.

[118] https://www.canada.ca/en/health-canada/services/chemical-substances/substance-groupings-initiative.html

the substances (e.g. ARN simple vanadium compounds, (ECHA, 2021); ARN simple lithium compounds, (ECHA, 2022a)). For example, cobalt- and selenium-containing substances were respectively grouped for risk assessment under Canada's CMP based on a 'moiety' approach. In the case of selenium, this corresponded to considering selenium in all oxidation states (selenate, selenite, elemental, selenide), organic selenium, and all forms of selenium that could contribute to cumulative loadings, exposure, and effects of selenium (ECCC, 2017).

When scientifically justified and after the grouping is done, read-across can be conducted within a group if there is sufficient supporting information and source substance data to predict hazard data for substances where no substance-specific information is available. The ECHA Read-Across Assessment Framework (RAAF) states that the read-across hypothesis for metals may often rely on the same metal ion being (bio)available (falling under the hypothesis '(bio)transformation to common compound(s)') (ECHA, 2017b).

While the assumption of basing the grouping on the extent of release of a 'common ion/complex' seems reasonable and straightforward for most inorganic metal substances, the selection and justification of the substances for which a grouping approach is relevant needs to be done with care and reported in a transparent and reproducible way.

This guidance is aimed at facilitating this process and is based on the learnings and experience with metals grouping in regulatory forums[119]. This section intends to supplement the general guidance in the previous chapters with issues specific to metals and inorganic compounds, using a weight of evidence approach to grouping of metal substances, illustrated by some examples.

It should be noted that this section focuses specifically on metals and inorganic metal compounds (and some organic metal compounds like metal salts of some organic acids). For more guidance on organometals[120] it is proposed to refer to the OECD N°212 Series on Testing and Assessment (OECD, 2017c). Also, this section does not address nanomaterials, which are discussed in Section 6.8.

6.7.2. Main hypothesis and limitations

The key hypothesis underlying the grouping of inorganic (metal) substances is that their (eco)toxicological properties are likely to be similar or follow a similar pattern if they have similar release rates or generate similar releases of the same common (metal) moiety under comparable exposure circumstances. It is the release of that (metal) moiety (including a hydrated metal ion or a redox form of the ion) from the metal containing material and its presence at target sites that is associated with (eco)toxicity (Goyer, 1996). Hence the "bioavailability" of that common (metal) moiety is what determines the hazard and potency of the effects being assessed.

Note other mechanisms of toxicity that are independent of metal bioavailability may exist. For example, after inhalation of poorly soluble particles, local lung toxicity in rodents can be observed, independent of the metal ion release. These effects are specific to a route of exposure (inhalation), particle size (respirable materials), and level and duration of exposure (repeated exposures to concentrations at which particle clearance is impaired), particularly in rats. Respirable TiO_2 and carbon black are examples of Poorly

[119] Examples include: Evaluation of EU REACH dossiers and Testing Proposals under EU REACH, ECHA-Member States workshop on metals read-across in 2012, the Metals and Inorganics Sectoral Approach (cooperative programme set up by ECHA and Eurometaux (European non-ferrous metal association)/metal sector to improve the quality and compliance of the REACH Registrations (2018-2022, https://www.reach-metals.eu/metals-and-inorganics-sectoral-approach-misa), NICNAS presentation at ICMM workshop (Philippines) AIMP: Meeting Document Database - Search Results (apec.org), IARC monographs on e.g. nickel and nickel compounds, cadmium and cadmium compounds etc.).

[120] Organometals (OM) include all compounds identified as coordination complexes where the metal or metalloid has covalent-character bonds with oxygen, nitrogen, sulphur and/or phosphorus belonging to an organic moiety.

Soluble, Low Toxicity Particles (PSLTs), for which toxicity (and even carcinogenicity) may be secondary to impaired clearance (e.g. (Driscoll, 2022)). Respirable crystalline silica is an example of a particle that can exert toxicity, independently of metal ion release and at concentrations below those that impaired particle clearance.

When assessing the validity and reliability of the main grouping or read-across hypothesis for a specific group of metals and/or inorganic substances, the following factors may need to be considered:

- Chemical structure
- Physical form and properties
- Speciation
- Crystal/amorphous structure
- Counter-ions and other metal ions

These factors need to be assessed when performing grouping and read-across as they may affect release, uptake of the common (metal) moiety as well as the toxicity that is associated with the non-common moiety ('counter-ion').

These factors, some of which may primarily affect human health or environmental endpoints, are identified in guidance documents such as the "*ECHA Guidance on information requirements and chemical safety assessment Chapter R.6: QSARs and grouping of chemicals"* (ECHA, 2008a), the ECHA RAAF (ECHA, 2017b) and the previous version of this document "OECD Guidance on Grouping of Chemicals, Second Edition" (OECD, 2017d):

- Chemical structure: Identification of a chemical substructure common to the source and target substances is a key aspect of grouping and read-across for metal substances, with the metal (or complex) being the primary common structure, and other possible common structures of the substances being related to the chemical forms (e.g. sulphides, oxides, ionic salts).
- Physical form and properties: Toxicity of metal-containing substances can vary with the physical forms (massive, powder) and physical properties (e.g. particle size, surface properties, porosity), and therefore are important to consider for grouping and read-across. The influence of particle size and surface properties for the local or systemic bioavailability and/or potency of the metal moiety should be evaluated since these can influence the extent of solubility in aqueous and biological media, and associated bioavailability at target sites and potency of effects. For example, it is widely recognised that metals may occur in massive, powder, or nanoforms[121].
- In the environment, the release of metal ions from metal substances, and therefore the environmental toxicity, depends on the particle size and exposed surface area (UN GHS 10th revised edition, annex 9.7[122]). For example, massive forms of metals typically have several orders of magnitude lower ecotoxicity compared to metal powders of the same substance.
    - In the human body, the particle size of a substance can influence the deposition behaviour in the respiratory tract, its clearance, absorption, bioavailability, and toxicity. For instance, a

[121] For hazard identification purposes, the diameter of the particle defines the form (UN, 2013; EC, 2008; and ECHA, 2024b):

- Massive has a particle diameter ≥1 mm
- Powder has a particle diameter ≥100 nm <1 mm
- Nanoform has a particle diameter <100 nm

In some cases, the word 'bulk' is used to describe the non -nanos form (i.e. powder and massive, with particle diameters ≥ 100 nm)

[122] Globally Harmonised System of Classification and Labelling of Chemicals (GHS), Tenth revised version, 2023; GHS/Rev.10e (https://unece.org/).

metal substance that exerts respiratory tract effects after inhalation would only exhibit this hazard if the powder can be inhaled. Also, particles with either positive or negative surface charge may be taken up by lung epithelial cells much more effectively than oppositely charged or uncharged particles, depending on the type of metal-containing particles, which can affect their potency. Particle size and surface area can also influence the extent of solubility in aqueous and biological media, which can affect bioavailability at target sites and potency of effects. For the oral route of exposure, the effect of particle size is less important for accessing the gastrointestinal system for local or systemic toxicity, but the influence of particle size and surface properties remains a factor for consideration in grouping and read-across in terms of effects on solubility of the metal moiety.

- Speciation: For some metal substances, the speciation may be associated with differences in mechanism of action and/or a variation in (eco)toxicological properties. Speciation may include different valences of the metal, or different forms of the substances, e.g. hydrous versus anhydrous, different oxidation states. Many metals ions exist predominantly in a single valence after their release into environmental or biological media. However, some metal ions have several valences which need to be considered in a grouping strategy as they may be associated with varying physicochemical properties and associated (eco)toxicity. A classic example for the role of 'speciation is the difference in hazards seen between trivalent and hexavalent chromium compounds, or for metals like vanadium (with valence 0, elemental), tri-, tetra-, or pentavalent substances, some having a respiratory irritation potential, others not). An example for the role of 'complexation' is the difference in sensitising potential between $[Pt(II)Cl_4]^{2-}$ (highly potent sensitiser) and $[Pt(II)(NH_3)_4]^{2-}$ (not sensitising) compounds. In some cases, chemical species may be interconvertible; in other cases, there is little interconversion between the species. It should be noted that when grouping is based on metal moiety release, extrapolating from metal compounds to the metal in elemental form (i.e. valence 0), or vice versa, may require a case-by-case approach as the metal undergoes corrosion processes that are substantially different from the pure dissolution processes seen with compounds. During corrosion, which is defined as *"[…] chemical or electrochemical reaction between a material, usually a metal, and its environment that produces a deterioration of the material and its properties"* (ASM, 1987), the metal reacts with air (oxygen) and/or water and the metallic form is transformed to the ionic state. Corrosion is strongly dependent on factors like the electrochemical properties of the metal (e.g. as pure metal or in alloys), the surface properties, the composition of the medium or biological condition when the process occurs inside an organism.
- The crystal/amorphous structure of inorganic substances could influence their hazard profile and should be considered in the evaluation of the grouping and read-across. If there is reason to believe that the crystalline structure significantly influences the effects of the substance to be assessed, this will constitute an additional line of evidence. An example is silica, of which the respirable crystalline and non-crystalline forms in the workplace have different physicochemical properties and classification (e.g. for the endpoint carcinogenicity; see category synthetic amorphous silicas assessed within the OECD Cooperative Chemicals Assessment Programme[123]). Another relevant example is the inorganic pigments, with most of them being of a spinel[124] character. Individual atoms in their crystal lattice structure can be substituted within ranges without altering the chemical inertness of the spinel itself, thus rendering them not bioavailable. When these ranges are

[123] Silicon dioxide [CAS Nos 7631-86-9, 112945- 52-5, 112926-00-8]; Silicic acid, aluminum sodium salt [CAS No 1344-00-9]; Silicic acid, calcium salt [CAS No 1344-95-2].

[124] Spinel – The spinel structure is a crystallisation pattern for a multitude of minerals. Spinels can be synthesised with the common structural formula $[A_XB_2\text{-}X]O_4$, "A" being a divalent metal and "B" a tri- or tetravalent metal, via sintering, hydrothermal synthesis or the Verneuil process. Since various metals can crystallise in the spinel structure, countless modifications exist.

exceeded, however, a metal constituent may become readily leachable from the pigment matrix, thus subjecting the pigment of such a composition to be placed in a different group for read-across purposes.

- Counter-ions and "other metal ions": The assumption that in most cases the metal moiety is responsible for the common property or effect implies that the toxicity of the counter-ion or of other metal substances present in the compound is (largely) irrelevant in producing the effect(s) to be assessed. However, the counter-ion should be evaluated explicitly for its own toxicity potential to confirm that it is not toxic "in its own right" or is at least sufficiently less toxic than the "driving" metal moiety and does not contribute to the toxicity of the substance. The influence of the counter-ion should be checked for each relevant endpoint.

How to consider the factors listed above is further described in the next section (see Section 6.7.3).

6.7.3. In practice

As stressed above, the selection and justification of the substances for which a grouping approach is relevant needs to be transparent and reproducible. Grouping will require different types of data to be used in a WoE approach. Building a matrix with all available information can be a way to organise the data and provide a transparent understanding of the grouping logic that is followed. This could be done using a table (see example Table 16) or a narrative describing the various datasets available for the substances under consideration. It is proposed to build a matrix including all available relevant data such as the one below, specified by (eco)toxicological endpoint and route of exposure as needed, and where information is available. Such a matrix will show the degree of similarities and differences among the group members and allow definition of the most appropriate grouping, identification of the source substance(s) for each target substance/group, and transparent communication of the grouping justification. This example of a matrix highlights the important general types of information to include in the assessment. More detailed information for human health and environment are included below.

Table 16. Matrix compiling all lines of evidence with listing of examples of relevant types of data (other relevant types can be added in additional rows)

Endpoint:					
Route of exposure or environmental compartment:					
Type of data/lines of evidence	**Source substance 1**	**Source substance x**	**Target substance 1**	**Target substance 2**	**Target substance y**
Chemical composition					
Physicochemical: water solubility, metal ion release in other aqueous media or synthetic fluids					
Counter-ion toxicity					
Other relevant data, including toxicity					

A first step is to properly characterise the substances included (or used for) the grouping and read-across approach(es). Basic data for chemical composition, like chemical formula and identifiers (CAS and/or EC number(s)), are a prerequisite to properly identify the substances. These can also include MW or % metal

content (MW basis) and other analyses as appropriate. Information on possible impurities should be considered as well. This information should always be available when describing/justifying a grouping or read-across.

How to complete and consider the other lines of evidence are discussed separately for human health and environmental endpoints below.

As shown in the following, many examples exist for how grouping and read-across have been used for metal substances. Important considerations in the WoE for grouping and read-across for metal substances are relevant physicochemical and toxicological data, including the extent and/or rate of release of a common metal moiety in relevant media and information provided by *in vitro* tests (e.g. those that measure mutagenicity, reactive oxygen generation, DNA damage, etc.). Templates such as those provided below are helpful to present the data and provide justification for grouping and read-across of metal substances.

Grouping for human health endpoints

Grouping for human health requires building a matrix with data for different lines of evidence, such as data relating to the release and bioavailability of the metal ion/moiety, as well as data on factors that may affect the assumption of commonality between the different metal substances included in the group. Many health effects are, however, specific to certain route(s) of exposure and in these cases, grouping and read across should be based on the route-specific datasets. Table 17 lists the types of data that may be available for consideration, for systemic or local effects, in addition to consideration for the route of exposure.

Table 17. Example of matrix compiling all lines of evidence for human health endpoints with illustration of relevant types of data (other relevant types to be added in additional rows)

Endpoint (e.g. repeated dose toxicity):						
Route of exposure:						
Types of effect (systemic-local):						
Type of data*	**Source substance 1**	**…**	**Source substance x**	**Target substance 1**	**…**	**Target substance y**
Chemical structure						
Speciation/valence						
Crystallinity						
Water solubility						
Metal release in synthetic biological fluids						
Physical form, particle size, surface properties						
Toxicity of counter-ion[125] (effect on bioavailability/toxicity?)						
Mode/mechanism of action						
In vitro toxicity tests						
Toxicokinetics						
Acute toxicity						
Repeated dose toxicity						
Other toxicity endpoints						

Note: * Relevant data (e.g. NOEC, EC_{50}, EC_{10}) for each type of data should be entered for each substance, together with the proper reference

[125] A document summarising the data on a number of typical metal counter-ions (N=10) and in particular the information of toxicity is available here: https://www.reach-metals.eu/metals-and-inorganics-sectoral-approach-misa)

As it is the bioavailability of the metal moiety at target sites that will determine the occurrence and severity of the effects, ideally bioavailability data (from for instance *in vivo* toxicokinetic studies) are available and support the grouping. However, bioavailability data is rather scarce and requires animal testing. In the absence of such data, it can conservatively be assumed that the different metal substances considered for grouping have the same bioavailability as the most soluble substance(s) in the group. Other types of data can be used to assess assumed similarities or differences in bioavailability like physochemical data (e.g. water solubility), *in vitro* data (e.g. metal(loid) release in artificial physiological media, *in vitro* toxicity tests), or *in vivo* studies (e.g. acute or subchronic toxicity data) to support grouping and read-across.

Water solubility can be used as an indicator of relative metal moiety release and bioavailability (OECD, 2017e), and has been used for grouping some metal substances (like harmonised classifications of nickel compounds under EU CLP). However, the use of water solubility data only as surrogate for bioavailability data is associated with uncertainties that should be acknowledged. An example of an exception is $BaSO_4$, baryte, which is highly insoluble in water but exhibits remarkable clearance from lung tissue, almost like a soluble salt. Water solubility (in distilled water or in more complex media such as synthetic bodily fluids is driven by the solubility product of the anionic and cationic moieties. Such water solubility data do not reflect the influence of different pH or redox conditions, or the effect of various ligands in physiological fluids on metal dissolution, uptake, and toxicity.

Other contributing data useful for grouping is metal/metalloid releases in artificial physiological media. These data can be used to compare the extent of the metal ion/moiety release from two or more substances of the same metal (source and target chemicals). The fluids most relevant to the main routes of exposure include gastric and intestinal fluids (oral route), interstitial, alveolar and lysosomal fluids (inhalation route), and perspiration fluids (dermal route). While there are no internationally recognised metal release protocols for human health, there are several ones that have been accepted by regulatory authorities in various jurisdictions. For example, a protocol using synthetic perspiration fluid (EN 1811) is used in EU for the classification of Ni-containing alloys as skin sensitisers.

When evaluating metal/metalloid releases in artificial physiological media, the relative metal/metalloid release (%) is the ratio of the measured value of the metal ion (e.g. expressed as µg metal released/g sample, or % of metal content released) of the target substance compared to those from one or more source substances containing the same metal, in the same fluid, and the same time point. This relative release can be used with other data in the WoE approach to group metal substances that release similar amounts of metal ion in a particular physiological medium to determine if target and reference substances can be predicted to have similar or different effects. This is further illustrated below.

It is well understood that bioavailability *in vivo* is a dynamic process, and an *in vitro* metal release test cannot quantitatively predict *in vivo* bioavailability. Indeed, in biological systems, the dynamics are different from the static conditions encountered with *in vitro* metal release testing. Metal release tests currently cannot mimic all key components of *in vivo* bioavailability (like competitive inhibition of uptake, transport mechanisms, absorption, metabolism, or interactions). However, the amount of released metal “available for absorption” may be measured using *in vitro* methods and the outcomes used to support grouping and read-across as one line of evidence. This was shown in a study by (Heim et al., 2020) that investigated nickel and cobalt ion release from the metals and several alloys in synthetic gastric, interstitial, and lysosomal lung fluids. Other studies by (Danzeisen et al., 2020) and (Verougstraete et al., 2022) have also measured metal release of cobalt substances in synthetic body fluids.

The possible contribution of other moieties in the metal substances (e.g. counter-ions, other metals, etc.) shall be considered as well in the grouping and read-across process. More specifically, some data should be collected to demonstrate if and how the counter-ion may affect the bioavailability of the metal moiety and if it is not contributing to the observed effects “in its own right.” In other words, this data confirm that the counter-ion(s) and impurities are less toxic than the “driving” toxicity from the metal moiety. An example is provided below for the category of ‘inorganic molybdenum substances’, where the key moiety is an

oxoanion (MoO_4^{2-}) and the counter-ion is sodium, calcium, or ammonium. In the case of sodium molybdate (Na_2MoO_4), the substance-specific Derived No Effect Level (DNEL) is 7.3 mg Na_2MoO_4/kg(bw)/d. When exposed to sodium molybdate at the Mo-based-DNEL level, the person would be exposed to ca. 1.63 mg Na/kg(bw)/d. For a default 70-kg person, this would be ca. 114 mg Na per day. Based on the dietary reference values proposed by EFSA in 2019, an intake of up to 2 g Na per day (2000 mg/d) is considered as safe for adults. Thus, a 70 kg person exposed to sodium molybdate at the DNEL would ingest less than 6% of the safe dose of sodium (114/2000= 5.7%). In this case sodium does not contribute significantly to the systemic toxicity of sodium molybdate and it can be considered justified to base the hazard assessment for sodium molybdate only on the molybdenum moiety.

Other data that are not specific to metal substances are also helpful to consider in the WoE approach to grouping. Toxicokinetic studies can be used to demonstrate similarities in bioavailability (e.g. by comparing target tissue levels of animals given the same dose of different forms of a metal). MOA data can provide additional insights based on what, where, and when toxicity occurs for the different substances.

Finally, the available toxicity data itself are important to compare and confirm if the similarities seen in the above parameters are consistent with the type and degree of toxic effects, *in vitro* or *in vivo*. Typically, acute or (sub)chronic toxicity studies can also be used for grouping to compare substance toxicity profiles and verify consistencies and inconsistencies in bioavailability and toxicity. Other types of relevant standardised toxicity studies could also be used in the WoE approach.

For the initial grouping as well as for any subsequent read-across, all available and relevant toxicity data should be included in the matrix. The data of the source substances should be properly included (including information like threshold values, target tissues and study reference), and the target substances be mentioned. This way, it's clear for the target audience what the scientific basis (and its strength) is behind the grouping and read-across, and what data are being read across to other substances.

Grouping for environmental endpoints

The grouping assessment of inorganic metal substances for environmental endpoints follow the same guiding principles that (eco)toxicological properties will follow a similar pattern if they have similar release rates in the environment or generate the same metal moiety under comparable environmental exposures. When evaluating the metal releases in the environment for grouping and read-across, specific approaches have been developed to measure and estimate the factors affecting metal releases in the environment, like speciation and evaluation of counterions and other metals. These topics are discussed below.

For most metal-containing substances, it is the bioavailable free metal ion that is, to varying degrees, liberated in aqueous exposures and serves as the common thread of structural similarity as well as the moiety of toxicological concern to base the grouping (Adams et al., 2020). The free metal ion (or the metal complex) is generally considered the ecotoxicologically relevant species, as the anions are released at levels typically far below their toxic thresholds and contribute a negligible amount to the overall toxicity. While the free metal cation is typically the moiety of interest, for some substances, it is rather the dissociation into a similar metal complex that justifies grouping.

The ECHA RAAF (ECHA 2017b) notes that the grouping and read-across to the bioavailable free metal ion can be assessed based on the rate and extent to which metals and sparingly soluble metal compounds can produce soluble ionic and other metal-bearing species in aqueous media. This can be tested experimentally using the Transformation/Dissolution Protocol (T/D P) (OECD Series on Testing and Assessment, number 29 – OECD, 2001b). The T/D P allows for direct measurement of the dissolution of metal-bearing species from metals and sparingly soluble metal compounds, providing support for grouping and read-across based on "transformation to common compounds".

The grouping of metal substances may be considered appropriate when:

a. Source metal substances transform to common compounds (e.g. free metal ion), with similar ecotoxicity patterns;
b. ecotoxicity relies entirely, or primarily, on the concentration of the free metal ion (or the metal complex) and the ecotoxicological contribution of the counter-ion(s) and impurities (if present) is negligible; and
c. source substances follow a regular pattern of solubility and speciation which can be determined by transformation/dissolution testing.

The following sections detail methodologies for confirming these criteria. It is proposed to collate all these lines of evidence in a table such as the following (Table 18).

Table 18. Matrix compiling all lines of evidence for the environmental compartment with listing of some relevant types of data (other relevant types to be added in additional rows)

Environmental compartment:				
Type of data*	**Source substance**	**Target substance 1**	**Target substance 2**	**Target substance 3**
Chemical structure				
Partitioning coefficient				
Oxidation state(s)				
Water solubility				
Common moiety release in relevant environmental media				
Particle size and surface area				
Degradation				
Toxicity of counter-ion[126] (effect on bioavailability/toxicity?)				
Bioaccumulation				
…				
Aquatic toxicity**				
Terrestrial toxicity				
Sediment toxicity				
…				
Toxicity modifying factors				
MOA				
Other toxicity				

* Relevant data (e.g. NOEC, EC_{50}, EC_{10}) for type of data should be entered for each substance

** A specification per trophic level (algae, invertebrates, fish, plants…) can be given

Solubility and dissociation

As noted above, it is widely acknowledged that the bioavailability, and hence the toxicity, of undissociated metal compounds/complexes is lower than that of the dissolved metal ions/complexes (further called 'moiety of ecological concern') (Adams et al., 2020). Measurement of dissolution in the T/D P test indicates the rate and extent of production of dissolved metal ions/ecotoxic moieties from metals or sparingly soluble metal substances. The rate and extent of release of these moieties is evaluated in standardised aquatic toxicity medium (pH range 6-9) at defined mass loadings of the metal or sparingly soluble metal substance and within defined timeframes. Standard ecotoxicity testing protocols (e.g. (OECD, 2004e) typically utilise soluble metal substances in order to maximise the ecotoxic moieties to elicit a high degree of ecotoxicity.

[126] A document summarising the data on a number of typical metal counterions (N=10) and in particular the information of toxicity is available here: https://www.reach-metals.eu/metals-and-inorganics-sectoral-approach-misa

Results of ecotoxicity testing with soluble metal substances can be used to derive thresholds, such as NOECs, EC_{50}s, EC_{10}s, Predicted No-Effect Concentrations (PNECs), or Ecotoxicity Reference Values (ERVs). These thresholds are protective for different trophic levels for various environmental compartments. The thresholds are then compared with the transformation/dissolution data collected using the T/D P to further identify how a substance will behave within the environment, assign appropriate hazard classification categories, and provide support for grouping and read-across based on transformation to common compounds. Since the T/D P is designed to be conservative, if one applies the T/D P for the environmental assessment by comparing the measured values of released ions with ERVs, this might lead to an environmental hazard classification although the substance when tested *in vivo* would not be classified.

As an example, grouping of copper substances for environmental hazard assessment is considered. Members of the group include the following: copper (Cu), copper dihydroxide ($Cu(OH)_2$), copper oxide (CuO), copper sulphate ($CuSO_4$), copper thiocyanate (CuSCN), copper(II) carbonate-copper(II) hydroxide ($CuCO_3.Cu(OH)_2$) , dicopper oxide Cu_2O), dicopper chloride trihydroxide ($Cu_2(OH)_3Cl$), tetracopper hexahydroxide sulphate ($Cu_4(OH)_6SO_4$), and bordeaux mixture (reaction products of copper sulphate with calcium dihydroxide). The read-across hypothesis is based on a category approach and on “transformation to common compounds”: due to aquatic speciation and transformation-dissolution processes, the environmental effects of these substances are driven by transformation to a common toxic moiety, the dissolved cupric ion Cu^{2+}. Transformation/dissolution data show that all these substances release copper ions into aquatic media to varying degrees, ranging from fully soluble substances such as copper sulphate to very low release from copper in its massive form (0.3% or less under all tested conditions). The available data on the environmental effects of copper have been obtained using fully soluble copper compounds – typically copper dichloride, copper dinitrate, or copper sulphate. The effects data are expressed based on dissolved or total Cu concentrations. The ERV is therefore expressed as dissolved Cu concentration.

- Copper is a metal and is therefore insoluble. To derive the environmental hazard classification for copper according to the UN GHS, the ERV was compared to the available T/D data. Both are expressed as dissolved copper.
- All copper compounds were considered as readily soluble due to the high toxicity of dissolved copper. To derive the environmental hazard classification for copper compounds according to the UN GHS, the ERV (expressed as dissolved copper) was read across to each compound by considering its copper content, i.e. by applying a correction for the MW of each compound. The resulting ERV of the compound was then compared to the classification criteria.

Ecotoxicity of the metal-ion and counter-ion

The ecotoxicity of the moiety of ecological concern and the counter-ions in a substance should be considered in a read-across assessment. The typical hypothesis that the toxic contribution of the counter-ion is negligible when compared to the toxicity of the moiety of ecological concern. To confirm this hypothesis, available ecotoxicology data can be used to define and compare the toxicity of the metal ion and counterion. As with many chemicals, studies often report a high degree of variability in toxicity across different species, endpoints, and test durations likely as a result of differences in bioavailability under different test conditions and resulting in different toxic mechanisms, so it is important to consider the variability across an ecotoxicity dataset.

Several peer-reviewed publications have investigated the acute and chronic toxicity of typical anions towards algae, invertebrates, or fish. Examples are the studies of (Elphick et al., 2011; Erickson et al., 2017; Mount et al.,1997; 2016; Simmons, 2012). These authors consistently conclude that acute and chronic toxic thresholds of typical anions are in the mg/L-range, for some even in the g/L-range, which is often orders of magnitude higher than those of the applicable ecotoxic moieties. Additionally, the low toxicity of typical anions is supported by the information that is provided in ECHA’s EU REACH registration

dossiers for simple salts like sulphates, chlorides, carbonates or nitrates paired with common inorganic cations such as calcium, potassium, magnesium, and sodium[127]. The available literature is often substantially more limited for sediment-dwelling organisms and terrestrial species than data that has been generated for the aquatic compartment however, toxicity threshold values for metals to terrestrial species and sediment-dwelling organisms are generally reported at levels approximately 4 orders of magnitude lower (i.e. more toxic) than the toxicity threshold values pertaining to the relevant counterions (Burton, 2010; Besser et al., 2013).

It is important to note that metal ions and counter-ions occur naturally in the environment, but their concentrations vary greatly depending on the site-specific geological, climatological and/or physicochemical conditions. In freshwater, for example, many of the typical counter-ions in metal compounds are found in concentrations in the order of mg/L while background concentrations for many ions-of-concern are generally found to be in the µg/L range (or even lower). Although the intrinsic properties of a substance are independent of the exposure scenario, it may be useful to understand the exposure chemistry when interpreting toxicity results or validating a read-across scenario to ensure that counter-ion concentrations do not exceed toxicity threshold values. In other words, an exposure analysis may be useful to validate the assumption that the toxicity of the counter-ion is negligible. To this end, when a grouping is proposed, it may be stated if there are boundaries associated with it.

Calculating ecotoxicity of the metal ion and counter-ions

In order to estimate the contribution of the counter-ion to the observed effect, the following approach is proposed. Assuming additivity of the toxicity of the moiety of ecological concern and the counter-ion, the corresponding Toxic Units ('TU') for the metal and the counter-ion are calculated as:

Box 6.4. Equation 4. Calculation of Toxic Units (TU)

$$TUi = \frac{[\,]i}{TTi}$$

Where:

- TUi = Toxic Unit for ion i
- [ ]i = concentration of ion i
- TTi = Toxic Threshold (like EC_x, NOEC, etc.) for ion i

The $TU_{counter\text{-}ion}$ needs to be considered with care, especially for the soil and sediment compartment, as reliable soil and sediment toxicity data are often limited (or even completely lacking). In these cases, metal toxicity data obtained using test methods that account for or minimise counter-ion effects (e.g. leaching, ageing, or salt control treatments) may provide useful additional context.

The effect of the added counter-ion can be assumed negligible for the most toxic metals. In these cases, the $TU_{metal\text{-}ion} >> TU_{counter\text{-}ion}$ at concentrations where effects occur, and the contribution of the added counter-ions to the observed toxic response can be ignored. In general, the effect of the added counter-ion is assumed to be negligible if:

$$\frac{\mathrm{TU(metal-ion)}}{\mathrm{TU(counter-ion)}} > 10$$

In this case, the observed effect can be solely related to the metal ion.

[127] https://chem.echa.europa.eu/

In contrast, for metals with lower toxicity, $TU_{metal\text{-}ion}$ can be comparable to or lower than $TU_{counter\text{-}ion}$ at concentrations where effects occur. In this case, the effect of the added counter-ions to the observed toxicity should be investigated and addressed in the assessment. In general, a more in-depth investigation is required if:

$$\frac{\mathrm{TU(metal-ion)}}{\mathrm{TU(counter-ion)}} < 10$$

In this case, the added counter-ion could significantly contribute to the observed effects, and further investigation of potential counter-ion effects is required.

Table 19 presents a framework for grouping substances in the environment based on ECHA's RAAF (ECHA, 2017b).

Table 19. Framework for grouping substances in the environment, based on (ECHA, 2017b)

Type of assessment		Source substance 1	…	Source substance x	Target substance 1	…	Target substance y
Substance characterisation							
Link of structural similarity and differences with proposed regular pattern		*[Following the categorical approach, describe the hypothesis supporting the approach and the structural similarities]*					
Solubility/K_{sp}							
Formation of common (identical) compound		*[Describe the mechanism(s) or condition(s) under which the common structure is formed]*					
Consistency of effects data on matrix		*[Describe any limitations or how the collected data ensures common structural linkage]*					
Degradation of non-common compounds		*[Describe chemical fate of non-common compounds (i.e. counter-ions)]*					
Bioaccumulation and ecotoxicity potential of non-common compounds		*[Describe bioaccumulation or toxicity potential of non-common compounds (i.e. counter-ions)]*					
Toxicity Threshold Values* (e.g. ERV/PNEC, etc.)	Metal-ion	**Selected threshold values (e.g. EC10s, EC50s) should be consistent between meta- ion and counter-ion*					
	Counter-ion						
Exposure concentrations	Metal-ion						
	Counter-ion						
Toxic Units (= Concentration/Toxicity Threshold)	Metal-ion						
	Counter- ion						
Ratio of Metal TU: Counter-ion TU							
Reliability and adequacy of source studies		*[Describe data-quality criteria used in screening of sources. Only studies determined to have highly reliable results with methodology thoroughly reported should be considered.]*					

6.7.4. Section 6.7 Annex – Case studies

- These case studies are illustrative examples only, and their publication does not imply acceptance of the methodologies for regulatory purposes across OECD Member Countries.
- In addition, they should not be interpreted as reflecting official regulatory decisions made by OECD Member Countries.

Human Health Example

Cobalt sulphate ($CoSO_4$) and cobalt metal (Co) have been found to cause lung tumours in rats and mice following inhalation exposure (Bucher, 1998; Behl and Hooth, 2014). Based on these data, both substances have classifications in Annex VI to the EU CLP Regulation as carcinogenic Category 1B (presumed to have carcinogenic potential for humans). The classification criteria under the GHS and EU CLP are similar. The predominant MOA of cobalt-related lung cancer is local chronic inflammation, e.g. (Danzeisen et al., 2022). Mutagenic responses have not been detected in guideline-compliant studies with any cobalt substance e.g. (Kirkland et al., 2015; OECD, 2014b).

Approximately sixty cobalt substances (including Co metal and Co sulphate) are registered under EU REACH. To fulfil data requirement and classification obligations, grouping was applied for local respiratory tract effects (e.g. carcinogenicity and repeated toxicity) after inhalation[128]. Co metal and Co sulphate were used as source substances with positive findings in carcinogenicity assays. Tricobalt tetraoxide (Co_3O_4) was used as a source substance for which lack of carcinogenicity was predicted based on lack of bioavailability, lack of induction of biomarkers of oxidative stress, cytotoxicity and hypoxia, no persistent inflammation after acute inhalation exposure, and overall low toxicity profile.

For the grouping to be robust, a series of physicochemical, *in vitro*, and short-term *in vivo* studies were conducted by Cobalt Institute on source and target substances. This work has been described in a series of publications (Danzeisen et al., 2022; Verougstraete et al., 2022; van den Brule et al., 2022; Derr et al., 2022; Viegas et al., 2022; Burzlaff et al., 2022) where a tiered approach to grouping is described. Here the data for 3 source (Co metal, Co sulphate, and tricobalt tetraoxide) and 3 target (Co dihydroxide ($Co(OH)_2$, Co dichloride ($CoCl_2$), and Co sulphide(CoS)) cobalt substances are presented to illustrate the types of data that can be gathered and the overall rationale that can be applied for grouping.

Types of data gathered

As indicated above, *in vivo* genotoxicity has been excluded as a major MOA of cobalt-related cancer. The extent to which a Co-containing substance causes inflammatory or pre-inflammatory events can be measured (e.g. using *in vitro* markers of hypoxia, DNA damage, cytotoxicity), and these 'markers' can be used together with *in vivo* data in a WoE approach to group Co-containing substances based on their predicted potential to cause chronic inhalation effects. See Table 20 for a compilation of the data generated from source and target Co substances.

- Physicochemical studies

These tests include Co content, valence, structure, and crystallinity; water and lung fluid solubility (Verougstraete et al., 2022); and particle size characterisation. In this example, no toxicity from counter-ions is expected (justification provided in Table 20)

- *In vitro* studies

[128] As a note, the evaluation process under EU REACH is ongoing and the example should not be taken to imply acceptance of the approach in this regulatory context.

Cobalt is known to be a potent trigger of a 'hypoxia-like' response (in normoxic conditions) and local damage in cells. Hypoxia and cytotoxicity are contributors to sustained inflammation and the development of cancer. Upregulation of biomarkers for these effects were studied in 2 different cell systems.

- Co substances were tested in human lung cells (A549) for two markers: cytotoxicity (WST-1) and hypoxia (HIF-1alpha) (van den Brule et al., 2022).
- A gene reporter assay (ToxTracker®) using mouse embryonic stem cells was applied to source and target substances to generate data on: p53, protein damage, oxidative stress, DNA damage, and hypoxia (Derr et al., 2022).

- *In vivo* studies:

These studies addressed relative bioavailability as well as inflammatory effects after inhalation exposure to various Co substances.

- Toxicokinetic studies after oral and/or inhalation exposure provide data on relative oral bioavailability (Danzeisen et al., 2020) and persistence of Co-containing particles after inhalation.
- Acute toxicity studies (4h) provide LC_{50} data that can be used to compare overall acute toxicities of Co substances (Viegas et al., 2022). Persistent inflammation and/or upper respiratory tract reactivity after acute inhalation exposure was also examined. 'Persistent' inflammation was defined as histopathologically visible markers of inflammation (inflammatory oedema (perivascular), alveolar pulmonary oedema and pneumonia) present two weeks after the acute exposure. Upper respiratory tract findings included changes in the larynx leading to epithelial hyper- or metaplasia, with scoring mirroring the system used for persistent inflammation (see Table 20)
- Four-week repeated dose inhalation toxicity studies. These studies were performed with 2 source substances (Co sulphate and tricobalt tetraoxide). Investigations of haematology, histopathology, immunohistochemistry (8-OH-dG) and bronchoalveolar lavage (BAL; total cell count, differential cell count, β-glucuronidase, total protein, LDH, HIF-1α, IL-8, MCP-1) were undertaken (Viegas et al., 2022). Similar data from 2-week and/or 14-week studies already existed for cobalt metal and Co sulphate (Bucher, 1998; Behl and Hooth, 2014).

Results

- Physicochemical studies

Significant differences in solubility and Co^{2+} release (spanning about 3 orders of magnitude) among substances and lung fluids were found. The source substance Co_3O_4 had a significantly lower solubility than any of the target or other source Co substances, in all fluids.

- *In vitro* studies

Testing of a human cell line indicated significant differences in cytotoxicity and induction of hypoxia across the Co substances. There were clear distinctions between four substances (Co metal, Co sulphate, Co dichloride and Co dihydroxide) positive for both endpoints, and two substances (tricobalt tetraoxide and Co sulphide) that were negative, were noted. Similar results were found with ToxTracker®; it predicted secondary/indirect genotoxicity (due to cytotoxicity, oxidative stress, and upregulation of hypoxia) for Co metal, Co sulphate, Co dihloride and Co dihydroxide but not for Co_3O_4 or Co sulphide.

- *In vivo* studies:

Acute toxicity studies yielded LC_{50} values for source and target substances that ranged from <50 mg/m^3 for Co metal and Co dihydroxide to > 5,000 mg/m^3 for Co_3O_4 and Co sulphide.

Toxicokinetic studies via oral route showed the bioavailability of Co from Co sulphide and Co_3O_4 to be similar and only 0.08% of that of Co chloride (Danzeisen et al., 2020).

Persistence of inflammation and/or 'upper respiratory tract reactivity after acute exposure was detected after exposure to Co metal, Co sulphate and Co dihydroxide, with increasing trends and normalized severity scores of 30 to 50 or >50. By contrast, Co_3O_4 and Co sulphide did not show persistent inflammation, had decreasing trends, and overall severity scores <1.

Rationale for grouping

For this cobalt example, the physicochemical data was helpful but did not, by itself, allow a definitive grouping of the target substances with either Co metal-Co sulphate or Co_3O_4 source substances (e.g. all the target substances released several-fold more Co^{2+} ion than Co_3O_4 and would a priori be grouped with Co metal-sulphate). The *in vitro* biomarker tests with either human or mouse cells were more informative, and they distinguished between the more reactive substances (Co metal, Co sulphate, Co dichloride and Co dihydroxide) that increased biomarkers of oxidative stress and hypoxia and the non-reactive ones (tricobalt tetraoxide and Co sulphide). They also confirmed the lack of direct genotoxicity of the Co substances.

A preliminary grouping based on *in vitro* assays was confirmed by *in vivo* studies looking at the persistence of inflammation and/or 'upper respiratory tract reactivity after an acute exposure. These studies also showed significant differences in response between Co metal, Co sulphate and Co dihydroxide, (very reactive), and the other two Co substances (tricobalt tetraoxide and Co sulphide).

With regard to repeated dose studies, the lifetime exposure dataset is currently limited to Co metal and Co sulphate. However, acute inflammation, subchronic and chronic effects correlate for both substances. Based on this limited dataset, there is a clear indication that "persistent inflammation" or "upper respiratory tract meta- or hyperplasia" is predictive of repeated dose inhalation toxicity. The results from 28-day inhalation studies with 90-day recovery periods for Co sulphate and tricobalt tetraoxide resulted in very different LOAECs for inflammation (1 and 20 mg/m^3 respectively, substance concentration and 0.21 and 14 mg Co/m^3, respectively, Co concentration). As suspected, given its low solubility in water and lung fluids, and its overall low toxicity, the quality of the inflammatory response elicited by tricobalt tetraoxide resembles that of poorly soluble low toxicity particles (PSLTs).

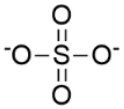
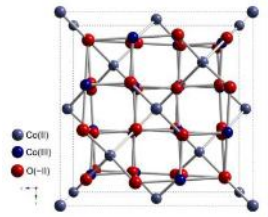
Based on a WoE integration of the data, Co dihydroxide and Co dichloride are grouped with Co metal-Co sulphate. These substances have (Co dichloride) or are predicted to have (Co dihydroxide) higher Co ion bioavailability *in vivo*, can elicit persistent inflammation in the respiratory tract after inhalation, and trigger indirect oxidative damage and hypoxia biomarkers *in vitro*. This is consistent with the proposed mode of action for the lung carcinogenicity of Co metal and Co sulphate. Repeated inhalation toxicity is also expected to follow a similar pattern for these Co substances.

By contrast, Co sulphide is grouped with tricobalt tetraoxide. The *in vivo* bioavailability of both substances is low, and they do not trigger persistent inflammation after acute exposure or induction of mode of action biomarkers *in vitro*. The repeated dose toxicity observed with Co_3O_4 in rats at much higher exposure levels in a 28-day study is characteristic of that of PSLTs. In cancer bioassays, PSLTs will not trigger lung carcinogenicity in rats, as long as particle clearance is not impaired, e.g. (Driscoll, 2022).

This example illustrates the usefulness of *in vitro* data on biomarkers and other endpoints (e.g. cytotoxicity) for preliminary grouping and the need to verify grouping with *in vivo* data, using a WoE approach.

Table 20. Matrix compiling all data sources available for the grouping of Co substances

The common compound formed by all substances within this group is the Co2+ cation. It is assumed that liberation of the common compound is mainly responsible for the local toxicity of many cobalt substances after inhalation. Particle effects of the less soluble Co compounds can also contribute to local toxicity.

Endpoints: carcinogenicity and repeated inhalation toxicity						
Route of exposure: inhalation						
Type of effects (systemic-local): local						
Type of data	**Source substance 1**	**Source substance 2**	**Source substance 3**	**Target substance 1**	**Target substance 2**	**Target substance 3**
	Co metal	**Co sulphate**	**Tricobalt tetraoxide**	**Co dihydroxide**	**Co dichloride**	**Co sulphide**
% Co (average Co content in substance)	99.9	21	72	62	25	63
Phys-chem: Chemical structure	Co	Co^{2+} 		Co^{2+} OH^- OH^-	Cl^- Cl^- Co^{2+}	
Phys-chem: Speciation/valence	Co^{2+}	Co^{2+}	Co^{2+}, Co^{3+} (mixed oxide). Aqueous only Co^{2+}	Co^{2+}	Co^{2+}	Co^{2+}
Phys-chem: Crystallinity	Cobalt is a chemical element and transition metal appearing in the fourth period of the periodic table between iron and nickel (group 9 of the d-block). It has an electron configuration of [Ar]	Chemical compound consisting of related numbers of cations (monoatomic or polyatomic species having one or more permanent elementary charges of the proton) and anions	Tricobalt tetraoxide/Cobalt(II,III) oxide is an inorganic compound with the formula Co_3O_4. It is a black antiferromagnetic solid. Co_3O_4 adopts the normal spinel structure, with Co(II) in tetrahedral interstices and Co(III) in the octahedral interstices of the cubic close-packed lattice of oxide anions (space group Fd3m, No. 227).	Cobalt(II) hydroxide or cobaltous hydroxide is the inorganic compound with the formula $Co(OH)_2$. Cobalt(II) hydroxide precipitates as a solid when an alkali metal hydroxide is added to an aqueous solution of Co^{2+} salt. Cobalt(II) hydroxide has the brucite crystal	Chemical compound consisting of related numbers of cations (monoatomic or polyatomic species having one or more permanent elementary	Cobalt(II) sulphide is an inorganic compound, crystallises in a hexagonal symmetry (space group P63/mmc, No 194)

Endpoints: carcinogenicity and repeated inhalation toxicity						
Route of exposure: inhalation						
Type of effects (systemic-local): local						
Type of data	**Source substance 1**	**Source substance 2**	**Source substance 3**	**Target substance 1**	**Target substance 2**	**Target substance 3**
	Co metal	**Co sulphate**	**Tricobalt tetraoxide**	**Co dihydroxide**	**Co dichloride**	**Co sulphide**
	3d7 4s2	(monoatomic or polyatomic species having one or more permanent elementary charges of the electron) so that the product is electrically neutral (without a net charge) (IUPAC 2005, 2014)		structure (rhombohedral, space group P63mc, No 186)	charges of the proton) and anions (monoatomic or polyatomic species having one or more permanent elementary charges of the electron) so that the product is electrically neutral (without a net charge) (IUPAC 2005, 2014)	
Phys-chem: Water solubility (20°C)	2.94 mg/L	376,700 mg/L	1.62 mg/L	2.3 mg/L	583,600 mg/L	15.2 mg/L
Phys-chem: Metal release in synthetic lung fluids[a]	% Co release after 72h: • Interstitial 4.0 • Alveolar 4.4 • Lysosomal 99 Co release (mg Co/g sample) after 72h: • Interstitial 40 • Alveolar 44 • Lysosomal 984	% Co release after 72h or 2h: • Interstitial 83 • Alveolar 50 • Lysosomal 79 Co release (mg Co/g sample) after 72h or *2h*: • Interstitial 173 • Alveolar 105 • Lysosomal 165	% Co release after 72h: • Interstitial 0.008 • Alveolar 0.02 • Lysosomal 2.2 Co release (mg Co/g sample) after 72h: • Interstitial 0.06 • Alveolar 0.02 • Lysosomal 18.0	% Co release after 72h or 5h: • Interstitial 1.6 • Alveolar 3.6 • Lysosomal 98 Co release (mg Co/g sample) after *5h* or *72h*: • Interstitial 10.2 • Alveolar 22.4 • Lysosomal 607	% Co content after 72h or *2h*: • Interstitial 46 • Alveolar 68 • Lysosomal 78 Co release (mg Co/g sample) after *72h* or *2h*:	% Co release after 72h: • Interstitial 0.3 • Alveolar 1.4 • Lysosomal 2.3 Co release (mg Co/g sample) after 72h: • Interstitial 2.0 • Alveolar 8.8 • Lysosomal 14.7

Endpoints: carcinogenicity and repeated inhalation toxicity						
Route of exposure: inhalation						
Type of effects (systemic-local): local						
Type of data	**Source substance 1**	**Source substance 2**	**Source substance 3**	**Target substance 1**	**Target substance 2**	**Target substance 3**
	Co metal	**Co sulphate**	**Tricobalt tetraoxide**	**Co dihydroxide**	**Co dichloride**	**Co sulphide**
					• Interstitial 196 • Alveolar 128 • Lysosomal 223	
Phys-chem: Particle size (dustiness test)	Powder, MMADs of airborne fraction: MMAD1 = 3.00 µm and MMAD2 = 25.66 µm GSD1 = 1.46 and GSD2 = 5.87	Representative sample is coarse material. MMAD of airborne fraction = 34.24µm. GSD = 1.56.	Powder, MMAD of airborne fraction= 18.23. GSD = 4.03.	Powder, MMAD of airborne fraction: MMAD1 = 2.30 µm and MMAD2 = 30.61 µm. GSD1 = 1.59 and GSD2 = 2.03.	Representative sample is coarse material. MMAD of airborne fraction = 33.07 µm. GSD = 1.43.	Powder, MMAD of airborne fraction = 31.47 µm. GSD = 1.62.
Toxicity of counter-ion[129] (effect on bioavailability/ toxicity?)	Not applicable	Sulphate [$SO_4^{(2-)}$] is widely distributed in nature as part of the sulfur cycle and is a highly regulated component of cells. Exposure to these ions in the amounts contributed by this cobalt substance will not trigger additional toxicity beyond that of the	Not applicable	Hydroxide [$OH^{(-)}$] as well as [$H^{(+)}$] and [$H_3O^{(+)}$], respectively are constituents and natural parts of water. Exposure to these ions in the amounts contributed by this cobalt substance will not trigger additional toxicity beyond that of the metal ion	Chloride [$Cl^{(-)}$] is widely distributed in nature and is a highly regulated component of cells (i.e. accountable in building a membrane potential). Exposure to these ions in the amounts contributed by	Not applicable

[129] A document summarising the data on a number of typical metal counterions (N=10) and in particular the information of toxicity is available here: https://www.reach-metals.eu/metals-and-inorganics-sectoral-approach-misa

Endpoints: carcinogenicity and repeated inhalation toxicity						
Route of exposure: inhalation						
Type of effects (systemic-local): local						
Type of data	**Source substance 1**	**Source substance 2**	**Source substance 3**	**Target substance 1**	**Target substance 2**	**Target substance 3**
	Co metal	**Co sulphate**	**Tricobalt tetraoxide**	**Co dihydroxide**	**Co dichloride**	**Co sulphide**
		metal ion			this cobalt substance will not trigger additional toxicity beyond that of the metal ion	
In vitro biomarkers (hypoxia and cytotoxicity) assay	*In vitro* stabilisation of hypoxia-inducible factor (HIF)-1alpha and induction of cytotoxicity: **Positive**	*In vitro* stabilisation of hypoxia-inducible factor (HIF)-1alpha and induction of cytotoxicity: **Positive**	*In vitro* stabilisation of hypoxia-inducible factor (HIF)-1alpha and induction of cytotoxicity: **Negative**	*In vitro* stabilisation of hypoxia-inducible factor (HIF)-1alpha and induction of cytotoxicity: **Positive**	*In vitro* stabilisation of hypoxia-inducible factor (HIF)-1alpha and induction of cytotoxicity: **Positive**	*In vitro* stabilisation of hypoxia-inducible factor (HIF)-1alpha and induction of cytotoxicity: **Negative**
In vitro gene reporter assay (p53, protein damage, oxidative stress, DNA damage, hypoxia) – ToxTracker® Assay indicate induction of genotoxicity, oxidative stress & upregulation of biomarkers of hypoxia	Activation of DNA damage markers: **Negative** Activation of oxidative stress response reporters: **Positive** Activation of several hypoxia target genes: **Positive**	Activation of DNA damage markers: **Negative** Activation of oxidative stress response reporters: **Positive** Activation of several hypoxia target genes: **Positive**	Activation of DNA damage markers: **Negative** Activation of oxidative stress response reporters: **Negative** Activation of several hypoxia target genes: **Negative**	Activation of DNA damage markers: **Negative** Activation of oxidative stress response reporters: **Positive** Activation of several hypoxia target genes: **Positive**	Activation of DNA damage markers: **Negative** Activation of oxidative stress response reporters: **Positive** Activation of several hypoxia target genes: **Positive**	Activation of DNA damage markers: **Negative** Activation of oxidative stress response reporters: **Negative** Activation of several hypoxia target genes: **Negative**
Toxicokinetic	Inhalation route: Tissue	Inhalation route: Tissue	Oral route: studies show >100-fold lower bioavailability	Not tested	Oral route: studies show	Oral route: studies show similar bioavailability as Co_3O_4 and >100-fold lower bioavailability than Co

Endpoints: carcinogenicity and repeated inhalation toxicity						
Route of exposure: inhalation						
Type of effects (systemic-local): local						
Type of data	**Source substance 1**	**Source substance 2**	**Source substance 3**	**Target substance 1**	**Target substance 2**	**Target substance 3**
	Co metal	**Co sulphate**	**Tricobalt tetraoxide**	**Co dihydroxide**	**Co dichloride**	**Co sulphide**
	concentrations of cobalt increased with increasing exposure concentration in all tissues examined.	concentrations of cobalt increased with increasing exposure concentration in all tissues examined.	compared to Co chloride		9.3% absolute bioavailability of Co from Co dichloride	chloride
Acute inhalation toxicity and *in vivo* persistent inflammation or upper respiratory tract meta- and hyperplasia, 2 weeks after acute exposure	Acute Inh. Toxicity LC_{50} < 50 mg/m³ 'Persistent inflammation' and/or 'upper respiratory tract reactivity score': **Positive (>50)** **Increasing Trend w time**	Acute Inh. Toxicity LC_{50} was not determined 'Persistent inflammation' and/or 'upper respiratory tract reactivity score': **Positive (≥30)** **Increasing Trend**	Acute Inh. Toxicity LC_{50} > 5,000 mg/m³ 'Persistent inflammation' and/or 'upper respiratory tract reactivity score': **Negative (<0.1)** **Decreasing Trend**	Acute Inh. Toxicity LC_{50} < 50 mg/m³ 'Persistent inflammation' and/or 'upper respiratory tract reactivity score': **Positive (≥ 50)** **Increasing Trend**	Not tested	Acute Inh. Toxicity LC_{50} > 5,000 mg/m³ 'Persistent inflammation' and/or 'upper respiratory tract reactivity score': **Negative (< 0.1)** **Decreasing Trend**
STOT RE inhalation	• 2-week study In the lung incidences of cytoplasmic vacuolization of bronchiolar epithelium significantly increased **Rat LOAEC 2.5 mg/m³** (2.5 mg Co/m³) In the lung alveolus histiocytic cellular	• 2-week study Adverse changes in the larynx **Rat LOAEC: 1 mg/m3** (0.21 mg Co/m3) • 16-day study Inflammation, necrosis in larynx, trachea, bronchioles and respiratory turbinates of the nose, degeneration	• 2-week study Lung weight statistical increase, significant increase in polymorpho-nuclear neutrophils (PMN) **Rat LOAEC: 48 mg/m³** (35 mg Co/m³) **Rat NOAEC: 12 mg/m³** (8.6 mg Co/m³) • 4-week study Neutrophil-driven, with the predominant hallmark of particle-laden macrophages and presence of free particles in the	Not tested	Not tested	Not tested

Endpoints: carcinogenicity and repeated inhalation toxicity						
Route of exposure: inhalation						
Type of effects (systemic-local): local						
Type of data	**Source substance 1**	**Source substance 2**	**Source substance 3**	**Target substance 1**	**Target substance 2**	**Target substance 3**
	Co metal	**Co sulphate**	**Tricobalt tetraoxide**	**Co dihydroxide**	**Co dichloride**	**Co sulphide**
	infiltration significantly increased **Rat LOAEC 5 mg/m³** (5 mg Co/m³) • 14-week study Mild to moderate lung chronic active inflammation, alveolar proteinosis, sometimes minimal hyperplasia of alveolar epithelium and minimal fibrosis of alveolar interstitium. Significant increase in erythrocytes, hemoglobin (upregulation of hypoxia) **Rat LOAEC: 0.63 mg/m³** (0.63 mg Co/m³) Minimal to mild bronchiolar epithelial hyperplasia **Rat LOAEC 1.25 mg/m³** (1.25 mg	of olfactory epithelium; hyperplasia in larynx; lung inflammation and histiocytic (macrophage) infiltration **Rat LOAEC: 50 mg/m³** (10.5 mg Co/m³) • 4-week study Lung inflammation, hypoxia upregulation, larynx squamous metaplasia **Rat LOAEC: 1 mg/m³** (0.21 mg Co/m³) • 13-week study Minimal to mild squamous metaplasia of the larynx. Damage to lung epithelia (e.g. necrosis, degradation). **Rat LOAEC: 0.3 mg/m³** (0.06 mg Co/m³)	alveolar space. **Rat LOAEC: 20 mg/m³** (14 mg Co/m³) **Rat NOAEC: 5 mg/m³** (3.6 mg Co/m³)			

Endpoints: carcinogenicity and repeated inhalation toxicity						
Route of exposure: inhalation						
Type of effects (systemic-local): local						
Type of data	**Source substance 1**	**Source substance 2**	**Source substance 3**	**Target substance 1**	**Target substance 2**	**Target substance 3**
	Co metal	**Co sulphate**	**Tricobalt tetraoxide**	**Co dihydroxide**	**Co dichloride**	**Co sulphide**
	Co/m³)	Chronic inflammation and squamous metaplasia in larynx, and histiocytic infiltrates in lung **Rat LOAEC: 1 mg/m³** (0.21 mg Co/m³) Increase in hemoglobin, erythrocytes **Rat LOAEC: 3 mg/m³** (0.63 mg Co/m³)				
Carcinogenicity inhalation	Carcinogenicity study Alveolar/bronchiolar adenoma or carcinoma in the lung of rats and mice; chronic inflammation of nose, larynx and lung	Carcinogenicity study Alveolar/bronchiolar adenoma or carcinoma in the lung of rats and mice; chronic inflammation of nose, larynx, and lung	Not tested	Not tested	Not tested	Not tested

a (Phys-chem: Metal release in synthetic lung fluids) - Italic values indicates that these substances only had 2-hour (or 5-hour) release rate data available.
MMAD: Mass Median Aerodynamic Diameter; GSD: geometric standard deviation; LOAEC: Low Observed Adverse Effects Concentration.

Environmental Example

Nickel and nickel compounds: the understanding of the physicochemical and ecotoxicological data for metals and counter-ions is essential to both understand the environmental fate and toxicological characteristics of the nickel compounds and to provide support for a grouping approach which would allow results obtained in tests conducted with soluble nickel salts to be considered applicable for all inorganic nickel compounds.

Solubility and dissociation

The evaluation of water solubility verifies that the free metal ion is a common product released by the chemically favourable dissociation of the parent substances.

Compound	Reaction	K_{sp} (solubility product constant)
Nickel dichloride	$NiCl_2 \rightarrow Ni^{2+} + 2Cl^-$	>486
Nickel dinitrate	$Ni(NO_3)_2 \rightarrow Ni^{2+} + 2NO_3^-$	8900
Nickel sulphate	$NiSO_4 \rightarrow Ni^{2+} + SO_4^{2-}$	>3.59

Relative ecotoxicity evaluation: metal-ion

The acute and chronic aquatic toxicity of nickel has been well studied in aquatic algae, invertebrates, and fish. As with many chemicals, the studies for nickel report a high degree of variability in toxicity across different species, endpoints, and test durations. Under the current protocols, nickel substances are considered together following the read-across approach, with the anionic component of the tested compound (if it is reported) being retained within the nickel databases. Table 21 displays values (NOEC/LOEC/EC_{10}) spanning from 0.04 – 0.47 mg/L across three taxa, considered as no- or low effect values. For nickel compounds, the chronic values are statistically similar to one another, suggesting that the free nickel ion dominates the chronic toxicity of all of the substances included in the proposed grouping.

Table 21. Chronic aquatic nickel toxicity for three taxa as a function of the counter-ion in the exposure. The chronic toxicity shown within this table includes reported NOECs, LOECs and EC_{10}s

Species	Median Aquatic Chronic Nickel Toxicity (mg Ni/L)		
	Chloride	Nitrate	Sulphate
Fish	0.47 n = 9	0.07 n = 1	0.11 n=3
Invertebrate	0.04 n = 29	0.23 n = 1	0.15 n = 8
Algae	0.04 n = 86	0.10 n = 3	0.16 n = 10

Relative ecotoxicity evaluation: counter-ion

Environmental fate and effects data evaluating the nickel anion's contribution to ecotoxicity is collected using a readily dissociable cationic component (e.g. the sodium salt) to counter-balance the ionic ratios while minimising any toxicity that would be the result of the cationic component. In this way, the anionic effects data can serve as surrogate for the anionic component of the corresponding nickel salt. Similarly, the metal ion can be represented by fate and toxicity data collected when the free nickel ion is counter-balanced with simple anionic components (e.g. chloride or nitrate salts).

Discrepancies between the relative toxicity thresholds of the metal ion and the counter-ion support the assumption of grouping of substances based on the free-metal ion being the primary driver of ecotoxicity. For example, the toxicity of the metal-ion is often in the order of µg/L while the toxicity of the corresponding counter-ions (e.g. chloride, nitrate, etc.) is in the order of mg/L.

Table 22. Key values for chemical safety assessments as reported in the substance EU REACH registration dossier for freshwater invertebrates[130]

Anion	Substance	Chronic Invertebrate EC_{10}/NOEC (mg/mL)	Mean EC_{10}/NOEC (mg/mL)
Chloride	$CaCl_2$	240	291.6
	$MgCl_2$	321	
	NaCl	314	
Nitrate	$Ca(NO_3)_2$	NA	608
	KNO_3	>245	
	$NaNO_3$	971	
Sulphate	$CaSO_4$	>100	604.5
	Na_2SO_4	1109	

Exposure scenario considerations

In freshwater, many naturally occurring anions are found in concentrations in the order of mg/L while background nickel concentrations are generally found to be in the µg/L range. In this example, the FOREGS Database (Salminen et al., 2005) is summarised in Table 23 to illustrate the chemical composition of more than 800 European natural waters. When an exposure scenario is considered, the difference in magnitude between background levels and toxicity threshold values (see Table 22 above) often illustrates that the counter-ion concentrations are not sufficient to elicit a toxic response relative to their toxicity threshold values.

Table 23. Ionic concentrations in European natural waters

Median ion concentrations in natural waters (data from FOREGS, mg/L)	
Nickel, Ni^{2+}	0.0019
Chloride, Cl^-	8.8
Nitrate, NO_3^-	2.8
Sulphate, SO_4^{2-}	16.1

[130] Data collected from echa.europa.eu/echa-chem on 9/22/2020.

Collect and summarise the data

Table 24 summarises quantitative and qualitative lines of evidence supporting the grouping and read-across for nickel ions in the environment. The ratio of the nickel TU to the counter-ion TU depends upon the robustness of measured concentrations and reported toxicity thresholds. Variability amongst these measurements may lead to increased uncertainty among the conclusions. For instance, although the chronic toxicity threshold for nickel sulphate does not differ significantly from the thresholds for either nickel dichloride or nickel dinitrate, the final ratio of nickel sulphate TU/counter-ion TU is significantly below the same ratio for both nickel dichloride and nickel dinitrate. This is due to the combination of lower nitrate concentrations in the environment and the lower toxicity threshold reported for nickel dichloride. Based on the recommendations made within this guidance, nickel dichloride and nickel dinitrate sufficiently exceed the TU ratio threshold of 10, while the TU ratio for nickel sulphate is less than 10 and may require additional assessment to justify in a read-across. The first step of performing this assessment should be to confirm the toxicity thresholds and reported concentrations used for the read-across.

Table 24. Summary of quantitative and qualitative lines of evidence supporting the grouping and read-across for nickel ions in the environment

		Nickel Dichloride	Nickel Dinitrate	Nickel Sulphate
Substance characterisation		$NiCl_2$ CAS: 7718-54-9 EC: 231-743-0	$Ni(NO_3)_2$ CAS: 13478-00-7 EC: 238-076-4	$NiSO_4*6H_2O$ CAS: 10101-97-0 EC: 232-104-9
Link of structural similarity and differences with proposed regular pattern		With respect to nickel substances, the free metal ion is considered the ecotoxicologically relevant species, whereas counter-ion species liberated upon dissolution are considered to have a negligible contribution to the overall ecotoxicity compared to the metal-ion. Thus, the category consists of inorganic metal substances for which the toxicity is governed by the free metal-ion.		
Solubility/K_{sp} (Solubility product constant)		>486	8900	>3.59
Formation of common (identical) compound		Nickel compounds release free-metal ions into the environment. The transport and bioavailability of the metal ion and associated counter-ions are determined by their solubility in environmental media (i.e. water, soils, sediments) which is driven by environmental conditions.		
Consistency of effects data on matrix		Data on environmental fate and effects was obtained using water-soluble metal salts which yield free metal ions in aqueous environments.		
Degradation of non-common compounds		Inorganic anions (e.g. Cl^-, NO_3^-, CO_3^{2-}, OH^-) are common in most natural environmental systems. They do not degrade but can form ionic bonds with other cations typically found in the environment.		
Bioaccumulation and ecotoxicity potential of non-common compounds		Inorganic anions (e.g. Cl^-, NO_3^-, CO_3^{2-}, OH^-) are common in most natural environmental systems and have a low bioaccumulation potential. Ecotoxicity potential of inorganic anions is evaluated below:		
Toxicity Threshold Values (e.g. ERV/PNEC, etc.)	Nickel Chronic Invertebrate EC_{10}/NOECs (mg/L)	0.04	0.23	0.15
	Counter-Ion Chronic Invertebrates EC_{10}/NOECs (mg/L)	292	608	604.5
Exposure Concentrations	Nickel (mg/L)	0.020		
	Counter-Ion	8.8	2.8	16.1
Toxic Units (= Concentration / Toxicity Threshold)	Nickel	0.5	0.09	0.13
	Counter-ion	0.030	0.005	0.027
Ratio of Nickel TU: Counter-ion TU		16.7	17.4	4.9
Reliability and adequacy of source studies		Only data and methodology considered high-quality were applied toward the read-across assessment. Specifically, reliable studies were defined as those having a Klimisch score of K1 or K2 within ECHA's EU REACH program.		

6.8. Grouping of nanomaterials

This section outlines considerations for the grouping and read-across of nanomaterials. Unless stated otherwise, the principles described in other chapters of this guidance document are generally also relevant and applicable to nanomaterials.

It provides context on the grouping of nanomaterials, explains specific characteristics that may influence their behavior and hazards -and thus are important to consider- and highlights known challenges. In line with the other chapters of this document, this section also mentions tools and approaches that can support grouping and read-across based on current scientific understanding. The use of these approaches depends on the context, the specific nanomaterials under consideration and applicable regulatory requirements.

General guidance notes for nanomaterials are also provided in Chapter 7.

6.8.1. Introduction

Nanomaterials[131] are a subset of chemicals distinguished by their nanoscale size. Although it is generally accepted that nanomaterials range from 1 nm to 100 nm (e.g. ISO 80004-2902 1:2023, and (EC, 2022)), national or regional definitions for nanomaterials can differ (Rasmussen et al., 2024). Nanomaterials can exhibit very different properties compared to larger (i.e. >100 nm) forms of the same substance, to the extent that chemical composition alone is not sufficient to predict hazard or risk.

Nanomaterials have a variety of physical (e.g. shape, surface characteristics) and chemical characteristics that influence their potential for toxic effects, fate and behaviour in the environment, uptake and toxicokinetics, and therefore influence considerations required for grouping. These nanomaterial characteristics impact for example (local) inflammatory responses, adsorption, persistence, carry-over effects, translocation and cellular uptake. In addition, a lack of standardised methods for characterising and assessing toxicity have impacted hazard assessment and grouping (Sellers et al., 2015), although in recent years, developments to standardise methods have improved.

Application of new methods and information have facilitated refinements in approaches for grouping nanomaterials (Worth et al., 2017; Giusti et al., 2019; OECD 2016c; OECD 2016d). Early approaches focused on associations between physicochemical characteristics and hazards (Kuempel, et al. 2012; BAuA, 2013; RCC, 2013a; RCC 2013b; Arts et al., 2015; Aschberger et al., 2019; Visser et al., 2022). These initial approaches for nanomaterial grouping proposed four broad groups defined by physicochemical properties and associated biological MOA: (1) soluble, (2) fibrous, (3) poorly-soluble low toxicity, and (4) high toxicity nanomaterials. Subsequent projects have also accounted for transformation of nanomaterials during different lifecycle stages (Stone et al., 2014; Wohlleben et al., 2019), including transformations within the environment and human body (e.g. changes in hydrophobicity, agglomeration, zeta potential[132] e.g. (Hund-Rinke et al. 2017)). Transformation can include modifications such as dissolution, agglomeration, or accumulation of a coating, which, like intrinsic physicochemical characteristics, can influence fate, behaviour, uptake, toxicokinetics (uptake, distribution, biopersistence) and hazard (including early and apical biological effects) (Oomen et al., 2015; Schwirn and Völker 2019).

The US National Institute for Occupational Safety & Health (NIOSH) evaluated the potential human health risk and the provision of the best available scientific information for risk management decision-making from

[131] The general term “nanomaterial” is used in this document, unless referring to the terms used in specific context, such as “nanoform” in the context of EU REACH (which is linked to the substance definition under this legislation). It should be noted however that generally these terms (such as nanomaterials, nanoparticles, nanoforms) may be used differently in different contexts and legislations.

[132] Zeta potential is the charge that develops at the interface between a solid surface and the liquid medium.

an occupational health perspective. NIOSH researchers proposed an evidence-based approach (Kuempel et al., 2012) that included comparative potency analyses amongst benchmark materials and new engineered nanomaterials within biological MOA classes. A literature-based data set focusing on acute rodent pulmonary inflammation was used in a proof-of-concept classification model (Drew et al., 2017) to group nanomaterials with similar hazard potencies and predict the potency group of a new nanomaterial based only on its physicochemical properties. In an extended literature-based dataset, Boots et al. (2021) found that a nanomaterial's chemical composition and form were the physicochemical properties most closely associated with its hazard potency. A NIOSH report (NIOSH, 2021) described the application of the (NIOSH, 2019) hazard banding framework to engineered nanomaterials.

Arts et al., (2015) described a decision tree tool for grouping nanomaterials focused on inhalation as the route of human exposure (DF4NanoGrouping project). A tiered approach to data collection was applied, in which Tier 1 included intrinsic physicochemical properties and identification of soluble nanomaterials. This was similar to the approach described in (Oberdörster et al., 2005). Solubility information permitted grouping of soluble nanomaterials with their molecular counterparts to facilitate read-across of inhalation hazards. Tier 2 compared system-dependent properties (e.g. agglomeration, reactivity, dissolution) against defined cut-off values derived from benchmark materials to generate three other grouping options: (1) passive, (2) biopersistent fibres, and (3) active materials in the lung. These groups are similar to the four that were proposed in earlier frameworks, e.g. (Hund-Rinke et al., 2017). Further division into subgroups was also suggested (Kuempel et al., 2012). Tier 3 utilised toxicological information (e.g. from short term *in vivo* inhalation studies) to corroborate the assignment of the nanomaterial to a group or subgroup. Applicability of the framework was subsequently addressed using different materials (carbonaceous, metal oxides and sulphates, amorphous silica, organic pigments) (Arts et al., 2016).

The ECETOC DF4NanoGrouping approach (Arts et al., 2015) was expanded by the German nanoGRAVUR project to applications beyond grouping for hazard assessment to enable grouping approaches for ecological risk and consumer risk, including aspects of lifecycle releases from composite nano-enabled products (Wohlleben et al., 2019; Arts et al., 2016). A second industry project led to the development of the ECETOC NanoApp (Janer et al., 2020; Janer et al., 2021) which attempts to address identification of 'sets of similar nanoforms' indicated in the EU REACH legislation.

The EU REACH legislation defines the 'set of similar nanoforms' as a group of nanoforms which have similar characteristics within clearly defined boundaries, which still allow to conclude that, for all the endpoints, the hazard, exposure and risk assessment of these nanoforms can be performed jointly. For example, nanoforms of titanium dioxide with the same crystal structure, shape and surface functionalisation may exist within a defined range of sizes but are not sufficiently different to alter the hazard to humans and the environment, i.e. a single hazard conclusion can be reached for this set of nanoforms within the defined range of sizes. This is different from grouping and read-across approaches, such as for data gap filling, which are endpoint-specific. Guidance regarding sets of nanoforms under EU REACH is available as Appendix for nanoforms applicable to the "Guidance on Registration and Substance Identification" (ECHA, 2022c). For grouping and read-across approaches, ECHA published guidance on grouping for nanomaterials as Appendix to the Guidance R.6-1 on Grouping (ECHA, 2019).

Another example of approaches is provided by the the EU-funded project Grouping, Read-Across, CharacterIsation and classificatiOn framework for regUlatory risk assessment of manufactured nanomaterials (GRACIOUS). The GRACIOUS effort built on previous efforts to develop a consensus framework to support the use of grouping and read-across of nanomaterials relevant to human health and environmental endpoints (Stone et al., 2020). The GRACIOUS framework provides examples on how to support grouping for nanomaterials, as included in Section 6.8.2.

Although a number of regulatory jurisdictions include provisions to use grouping approaches for the assessment of nanomaterials, specific approaches vary with legislation e.g. (Mech et al., 2019). Regulatory

requirements for grouping and read-across need to be taken into consideration for any grouping approach applied.

6.8.2. Development of a grouping and read-across approach for nanomaterials

Grouping hypothesis and identification of analogues

Before initiating the grouping, specific regulatory requirements must be considered as they may influence the regulatory acceptance of the grouping and read-across. As for conventional chemicals, the first steps for grouping and read-across for nanomaterials are problem formulation as in Step 0 of the analogue/category approach for chemicals (Sections 4.2.1, 5.2.1), determining data gaps to be filled by read-across as in Step 1 (Sections 4.2.2, 5.2.2), and consideration of whether the nanomaterial is a (potential) member of an existing category as in Step 2 (Sections 4.2.3, 5.2.3).

A grouping hypothesis can support the identification of adequate analogues substances/category members (Step 3, Sections 4.2.4, 5.2.4), as described below.

Category and analogue approaches are generally applicable for nanomaterials. In both approaches different group members could consist of different forms of the same substance (e.g. TiO_2 forms varying in size (including non-nanomaterial forms), shape, crystal structure or coating), or different substances (e.g. different metal oxides), although there may be differences regulatory authorities regarding "flexibility" among members of a group.[133] In order to justify grouping and read-across, in particular for regulatory purposes, it is important to demonstrate the similarity of the caused effects between the source and target material(s).

For the identification of potential analogues, compared to other types of substances, availability of (adequate) data may be limited for nanomaterials, and though databases exist for nanomaterial safety information, the data included may not be adequate for supporting grouping, considering that methods used to generate the data may not have been standardised (use of standard operational procedures) or validated yet. Moreover, there are no well-established tools available for identifying source analogues via a database**.** More FAIR data (findable, accessible, interoperable, reusable) is required to enhance possibilities to support grouping and prevent replication of data generating studies (Jeliazkova et al., 2021). Furthermore, (Q)SAR models for nanomaterials were so far seen as not sufficiently developed to support grouping and hazard assessment (Basei et al., 2019; Lamon et al., 2019a, 2019b). Further increasing knowledge on nanomaterial properties and MOAs (see for example (Attarilar et al., 2020)), and ongoing research to develop adequate (Q)SAR models for nanomaterials (some examples reviewed in (Li et al., 2022)) will contribute towards the use of (Q)SARs to support grouping and read-across of nanomaterials.

Grouping hypotheses for nanomaterials should be based on a combination of relevant physicochemical characteristics (see Section 6.8.3), route of exposure and a specific hazard endpoint (e.g. Murphy et al., 2023). As for other types of chemicals, the grouping and read-across is endpoint-specific. For each hazard endpoint, multiple types of evidence, each mapping to different parts of the MOA (or AOP) if known, should be considered. A systematic approach is suggested for gathering the evidence needed to assess the grouping hypothesis.

The grouping hypothesis should consider:

- Physicochemical characteristics (Loosli et al., 2022)
- Information on fate and behaviour in the environment and/or toxicokinetics
- Potential hazards informed by the MOA if possible

[133] For example, under EU REACH Regulation.

Similarity in chemical structure is not sufficient for grouping and read-across of nanomaterials. Instead, a range of physicochemical characteristics specifically relevant for nanomaterials need to be considered and generally includes intrinsic properties such as size, shape, composition, contaminants, crystallinity, or surface chemistry (e.g. coating or functionalisation), as well as extrinsic properties such as dissolution rate (in specific media) (Worth et al., 2017; Stone et al., 2020). Section 6.8.3 provides more detailed information on types of physicochemical characteristics to consider in the grouping and read-across of nanomaterials.

Intrinsic and extrinsic properties of the nanomaterials play a crucial role in their grouping. Intrinsic properties, such as chemical composition and structural characteristics, determine the fundamental nature of the materials, while extrinsic properties, like behaviour in various media, illustrate how materials interact with their surroundings. Together, intrinsic and extrinsic properties are essential for assessing the safety of nanomaterials (see also Section **Error! Reference source not found.**). Distinguishing between intrinsic and extrinsic properties, when used for grouping hypotheses, may be important because extrinsic properties like dissolution rate depend on the characteristics of the surrounding. Variations in dissolution rate between different media may provide challenges to grouping, but this can be mitigated by specifying the medium in which dissolution rate is most relevant and assessed (e.g. in lung lining fluid or in an artificial lysosomal fluid). It is important to note that, to effectively group nanomaterials, data for comparing the nanomaterials must have been obtained using consistent experimental protocols and the same composition of test media.

Any initial grouping hypothesis should be considered as draft and evidence generated in subsequent steps may be used to refine the hypothesis as appropriate, in an iterative process.

Data gathering

Once a draft of the grouping hypothesis is generated and analogues identified, available data on physicochemical properties, environmental fate parameter(s), ecotoxicological and toxicological effects should be gathered, which will also contribute to evaluating the grouping hypothesis.

The data gathering for nanomaterials can be considered to be equivalent to Step 4 which outlines data gathering for analogue substances (Section 4.2.5) or category members (Section 5.2.5). Similarly to Step 5 (Sections 4.2.6, 5.2.6), a data matrix should be constructed.

A possibility is to gather and organise data following examples from the GRACIOUS Framework aiming at supporting nanomaterial grouping. The Framework proposes the use of data of a range of *in vitro* methods covering for example particle reactivity, uptake into cells, cytotoxicity, and pro-inflammatory potential as supporting information to establish similarity (e.g. Verdon et al., 2022; Murphy et al., 2021; Braakhuis et al., 2021; Di Cristo et al., 2021; Di Cristo et al., 2022a). These published examples were used to systematically identify, collect, or generate physicochemical, fate, toxicokinetic and hazard data (e.g. using *in vitro* methods) needed to test the grouping hypothesis, however, it is important to note that these *in vitro* methods have not been validated for regulatory purposes. The comparability of data should be ensured by using the same test protocols for generating data for the substances to compare. Preference should generally be given to validated test methods, i.e. OECD test guidelines. The published examples were also informed by MOA/AOP information (Gerloff et al., 2017; Murphy et al., 2021). The GRACIOUS Framework is intended to guide data collection and data structuring, though the approach has not been reviewed by OECD.

The GRACIOUS Framework proposes a tiered testing (or data collection) strategy, related to specific grouping hypotheses. The choice of types of data to be collected can be driven by several factors such as the availability of the methodology, the variability of the data generated by a specific method, interference by the nanomaterials in a specific assay, or lack of availability of a method. For example, results from a lower tier method (e.g. *in vitro*, *in silico, in chemico* approaches) may be too variable to allow for a similarity assessment, prompting the user to move to a higher tier method which might have less variability and therefore allow for a more appropriate assessment of similarity. Such a situation was observed for

assessment of silica nanomaterial similarity in gut *in vitro* models (Di Cristo et al., 2022c), where a simple one cell type Caco-2 culture was too variable, while a 3D multilineage model provided less variation.

For any new data generated following a testing strategy, such as physicochemical characteristics, fate, toxicokinetic or hazard data, with the aim of strengthening the grouping hypothesis, the data should be generated using the same experimental method and composition of test media across all group members. Different methods may have different sensitivities or measure different parameters. Use of the same method allows for a comparison of similarity without introducing additional uncertainty. Existing uncertainties related to, for example, existing data generated with different types of methods, need to be taken into account in the assessment.

Adequacy and quality of data

Before assessing similarity, the adequacy and quality of the data and the basis of the comparison (including the source data) need to be considered (equivalent to Step 6 in Sections 4.2.7 and 5.2.7 - Availability and quality of data). The adequacy (e.g. for the endpoint considered) and quality of the data that can be accepted will depend on the specific purpose for which the grouping is being applied.

The assessment of data quality for nanomaterials has been discussed for example in (Comandella et al., 2020; Marquese Robinson et al., 2016; Hartmann et al., 2017; Lubinski et al., 2013; EFSA Scientific Committee 2021; Fernández-Cruz et al., 2018).

Compared with non-nanomaterial chemicals, there are relatively few standardised methods available for assessing physicochemical characteristics, fate, toxicokinetics, and hazards of nanomaterials, and therefore data to describe nanomaterials have uncertainties. Efforts by the OECD and other groups are addressing this current limitation in standardised test methods which will improve the uncertainties of similarity assessments related to the data in the future. The limitations in method availability are discussed further in Section 6.8.4.

Similarity assessment to confirm the group or refine the grouping hypothesis

This relates to Step 6 in Sections 4.2.7 and 5.2.7- Adequacy of the analogues/ category members.

When assessing similarity, a thorough and inclusive analysis of all data in a specific data set should be employed. As with other types of substances, a weight of evidence approach should be applied, usually by expert evaluation of the available data.

Methods and tools described in the next paragraphs might assist in the assessment of similarity. For example, (Basei et al., 2021) proposed a quantitative WoE method in the context of the GRACIOUS project. Moreover, several quantitative tools for assessing similarity have been described using nanomaterials as a case study (Jeliazkova et al., 2022a and 2022b; Tsiliki et al., 2022; Zabeo et al., 2022; Seleci et al., 2022; Ruggiero et al., 2022; Song et al., 2022). Assessment of similarity might be defined by cut-offs or thresholds (e.g. according to the World Health Organization (WHO) a fibre is described as *"a sample containing >0.1 % of inhalable particles being >5 µm in length, <3 µm in diameter, and with an aspect ratio of >3:1"*). However, care should be taken as not all applied cut-offs are widely available or agreed internationally for different parameters, thus they can be difficult to justify. To clearly define a cut-off value, nanomaterials that are systematically altered with respect to a single characteristic are required. This is difficult to achieve because such panels of nanomaterials are difficult to make and changing one parameter often influences another physicochemical characteristic.

To avoid the need for strict thresholds or cut-offs for all parameters, quantitative context-dependent similarity methods have been developed, e.g. (Park et al., 2018). Quantitative methods for assessing similarity reduce the need to rely on subjective expert judgement (Jeliazkova et al., 2022a) for specific parameters. Similarity can be assessed in a pairwise manner between pairs of nanomaterials for each

parameter or assessed using multidimensional models across all nanomaterials for all data sets simultaneously. For example, for a pairwise similarity assessment, data distributions (e.g. dose-response curves) can be reduced to a single value (e.g. LC_{50}, NOAEL) to simplify comparisons of similarity across the proposed group members, although other methods exist that prevent the loss of information inherent to such reduction (e.g. Ag Seleci et al., 2022). Quantitative similarity methods compare nanomaterials (and other group members) in a pairwise manner for each parameter descriptor considered in the data matrix. A number of algorithms for assessing similarity for a limited and non-representative set of nanomaterials and assays have been shown to be useful. Examples of algorithms include:

- A Bayesian model assessment to compare two sets of values using nested sampling (Tsiliki et al., 2022).
- An Arsinh-Ordered Weighted Average model (Arsinh-OWA) which applies the arsinh transformation to the distance between two nanoforms (Zabeo et al., 2022). The results are rescaled to the arsinh of a biologically relevant threshold before grouping using the OWA based distance.
- An x-fold comparison (Ruggiero et al., 2022; Cross et al., 2022), as used in the ECETOC NanoApp (Janer et al., 2020; Janer et al., 2021) for sets of similar nanoforms.
- Euclidean distance (Jeliazkova et al., 2022b; Cross et al., 2022; Di Battista et al., 2024), which commonly used as a distance metric.

The Bayesian and Arsinh-OWA methods were newly applied to nanomaterials. The x-fold, Bayesian and Arsinh-OWA distance algorithms were comparable in performance, with respect to the scoring of similarity between pairs of nanomaterials (Jeliazkova et al., 2022a). The Euclidean distance was also useful, but only following an appropriate data transformation. Whilst the x-fold method does not require standardisation of data, histograms produced are skewed. However, the advantage of the x-fold method is that it can be implemented without the need for programming expertise.

Ag Seleci et al., (2022) developed a novel method to compare dose-response curves to assess similarity, which required a four-step workflow on concentration-response curves, individual concentration and response ranges, and representative materials. The algorithm was found to be applicable for a range of abiotic and *in vitro* assays. Alternatives to reducing dose-response data to a single value include using Bayesian additive adaptive basis tensor product models previously developed for chemicals (Wheeler et al., 2019), but these methods have not been applied to nanomaterials.

For the example of the GRACIOUS Framework, multidimensional analysis methods which assess all data for all nanomaterials for all considered decision nodes have been assessed using hierarchical clustering approaches (Jeliazkova et al., 2022a). Whilst such methods may identify interesting patterns in data, the results do not always align with expert inference. Moreover, different multidimensional methods may give rise to different results (Jeliazkova et al., 2022a). Further work is required to improve the robustness of these outputs and their scientific relevance using additional data sets. Efforts to identify and address uncertainties due to factors such as statistical variability, data and measurement limitations, or biological differences (e.g. species-to-species extrapolation) are also needed.

Furthermore, as described in Sections 2.1.4 and 2.4.2, omics technologies such as transcriptomics, proteomics and metabolomics can be applied to group substances that induce similar bioactivity profiles. These technologies measure broad molecular responses, including for example perturbations in energy metabolism, oxidative stress and cellular signalling, which reflect underlying MOAs. Although primarily explored for the grouping of individual substances (Viant et al., 2024a), bioactivity-based grouping using omics data should be equally applicable to nanomaterials, with a range of options for introducing this data type as a line of evidence in a grouping hypothesis outlined in Table 1. This data-driven strategy can complement traditional structural grouping, including for poorly characterised nanomaterials, by quantifying bioactivity similarities statistically. For example, multi-omics studies have successfully clustered seven well-characterised nanomaterials based on their distinct metabolic and proteomic responses in a rat

alveolar macrophage cell line (NR8383), revealing correlations between agglomerate size, surface area and toxicity endpoints (Shaw et al., 2008).

Overall, similarity between source(s) and target(s) needs to be comprehensively evaluated, taking into account all relevant parameters, in order to confirm or refine the analogues/category members.

Assessment of the adequacy of the grouping and read-across approach

As for other types of chemicals, after data collection and/or generation, sufficient similarity should be demonstrated based on the gathered evidence for each parameter, and the formed (and possibly, refined during the iterative process) group should be confirmed as robust. Furthermore, the read-across hypothesis should be confirmed based on the available supporting information as well as the adequateness of the read-across approach. Data gap filling by read-across of nanomaterials requires similarity in a range of physicochemical characteristics, fate, behaviour, toxicokinetics and hazard (including bioactivity) data (Stone et al., 2020). This relates to Step 7 about assessing the adequacy of the analogue/category approach, Sections 4.2.8 and 5.2.8.

As for other types of chemicals (Section 2.3.2), new nanomaterials may be added to an already existing group. In case the data for the new nanomaterials indicate they fall on either the lower or upper end of an existing category, additional testing might be necessary to confirm that the nanomaterials belong to the group. As for chemicals, the concept of applicability domain, clearly defining the boundaries and inclusion/exclusion criteria (Section 2.3.2), is essential to assess whether a nanomaterial could be a member of a group, and to ensure the data is reliable for decision making.

Using the read-across approach for data gap filling

Once the grouping hypothesis has been evaluated and confirmed to be adequate and the grouping to be robust, the read-across for data gap filling can be performed.

If sufficient and adequate source data are available, the read-across can then be conducted, from the source(s) to the target nanomaterial(s). This step is equivalent to Step 8 in Sections 4.2.9 and 5.2.9.

To date, due to a lack of data and therefore availability of analogues to build a category, the analogue approach (see Chapter 4) has been used more frequently for nanomaterials than the category approach (as in Chapter 5). Examples of where the analogue approach has been used to fill data gaps include metal oxide nanoparticle induced hazards in lettuce (Song et al., 2022), multiwalled carbon nanotubes in lung (Murphy et al., 2022) and organic pigments in the lung (Jeliazkova et al., 2022b)

Documentation of the grouping and read-across approach, justifications and description of remaining uncertainties, in the context of the problem formulation, should be performed in analogy to Step 9 in Sections 4.2.10 and 5.2.10.

6.8.3. Properties and characteristics of nanomaterials relevant for hypothesis generation

Certain chemical legislations, such as EU REACH (Annex VI, Section 2.4), outline specific physicochemical parameters that must be considered for substance identification and characterisation of nanoforms and therefore should be explicitly reflected in any grouping approach. Additionally, going beyond the legislated physicochemical parameters to include other relevant ones that are related to the nanomaterial effects is generally encouraged. Several ongoing OECD projects aim to update the relevant OECD test guidelines to better assess these physicochemical parameters, addressing current gaps and therefore contributing to enhance the robustness of grouping strategies.

For nanomaterials, chemical similarity alone is not sufficient for the application of an analogue or category approach (see Section 6.8.2). In particular, a number of different, nanomaterial-specific aspects (see the

bullet points below) need to be considered to identify nanomaterials and evaluate their suitability as analogues for grouping and read-across (Worth et al., 2017). These physicochemical characteristics are not exhaustive or prescriptive but are intended to be flexible as needed for the group being developed and linked to specific hazard outcomes. The physicochemical characterisation data used to compare the group members should be generated using the same method across all group members.

It is worth noting that physicochemical characteristics are typically characterised by distributions (e.g. range of sizes, range of rates of dissolution) which can be visualised as a "cloud" or "sphere" with overlapping chemical and physical properties.

- **Chemical Identity of nanomaterials:** Different forms of the same chemical substance at the nano-scale differ in their intrinsic and extrinsic properties which could lead to different hazard profiles. Accordingly, hazard grouping should not solely rely on chemical composition and structure or on the CAS numbers (which are not unique).
- **Intrinsic and extrinsic characteristics of nanomaterials:** A range of intrinsic and extrinsic physicochemical characteristics need to be considered (Worth et al., 2017) as each may impact on their fate and behaviour in the environment, toxicokinetics, uptake, and hazard (Stone et al., 2020; Braakhuis et al., 2014). These characteristics include:
    - **Size (distribution):** Size will determine the behaviour and transport of nanomaterials in all environments (Williams et al., 2019). It will also influence entry into the body, for example, particle size determines deposition location in the lungs following inhalation (ICRP 1994). Particle size also impacts on uptake into cells and/or translocation across cell barriers. Nanomaterials entering the blood have been shown to distribute in a variety of organs (Kreyling 2014, 2017).
    - **Shape:** Nanomaterials can exhibit various shapes such as spherical, fibre-like, platelets and complex structures. The most well documented example of shape influencing particle toxicity is for high aspect ratio fibres. For example, asbestos is associated with lung fibrosis, lung cancer and mesothelioma (cancer in the pleural cavity) (Donaldson et al., 2010). Studies with high aspect ratio nanomaterials suggest that such hazards are relevant.
    - **Crystallinity:** Nanomaterials of the same substance can exist in various crystallinities which can influence the reactivity and toxicity. For non-nanomaterials such as alpha quartz (a crystalline form of silica), crystallinity was clearly associated with toxicity endpoints such as carcinogenicity, whereas amorphous silica particles were determined to be relatively benign (IARC, 1997). Differences in crystallinity of some nanomaterials have also been demonstrated to influence toxicity, such as TiO_2 nanomaterials which are available in rutile and anatase forms (Warheit et al., 2007).
    - **Surface properties:** As the surface of particles interacts with the environment and the human body, this has great potential to influence the hazard posed by nanomaterials. The surface properties of a nanomaterial result from the interaction of several interacting characteristics. Surface properties include for example:
        - Surface chemistry: Which is determined by chemical structure.
        - Surface reactivity: For example, a number of different nanomaterials have been shown to generate reactive oxygen species that can result in oxidative stress, inflammation or cell death (Oberdorster et al., 2007; Kessler et al., 2023; Seleci et al., 2022).
        - Surface area: Size and surface area are related, but the surface area can be further enhanced by a complex (e.g. rough or crenulated) surface structure. The surface area has been linked to hazard *in vitro* (Brown et al., 2001) and *in vivo* (Duffin et al.,

2002; Duffin et al., 2007) for nanomaterials with a very low dissolution rate (e.g. carbon black).

- Surface charge: As measured by zeta potential, can influence agglomeration, fate, and uptake.
- Surface hydrophobicity: Which might influence behaviours and compartmentalisation in hydrophobic and hydrophilic environments, including interactions with the cell surface and cell barriers.
- Coating: Coatings can modify the surface chemistry and therefore the surface properties. For example, addition of polyethylene glycol (PEG) to reduce interaction with cells enhances circulating time in the blood (Li et al., 2009).

- **Contaminants/impurities:** The presence of impurities should be assessed, in relation to the change or contribution to the mode of action of each individual nanomaterial. Such impurities may necessitate individual nanomaterials to be excluded from a presumed group.

- **Processes influencing the physicochemical characteristics of nanomaterials:**
  - **Agglomeration:** Agglomeration involves lower energy interactions between particles that can be broken by mixing or sonication. Agglomeration of nanomaterials occurs in many environmentally and physiologically relevant media. In the environment, agglomeration will influence the stability of the suspension in the water column and enhances settling and thus will change bioavailability in test systems which in turn will influence the hazard outcome, e.g. in ecotoxicological studies. Agglomerates may also be less mobile in soils and sediments (Williams et al., 2019). In the body, agglomeration may influence deposition in the lungs, uptake into cells and the ability of nanomaterials to cross cellular barriers. The availability of a reactive surface might also be influenced by agglomeration. For *in vitro* studies, agglomeration will also affect sedimentation and therefore the dose rate of nanomaterial delivery to the cells (Cohen et al., 2013; DeLoid et al., 2014) as well as the form of the nanomaterial coming into contact with the cells (Wick et al., 2007).
  - **Dissolution:** Consideration of dissolution allows the user to identify whether exposure at a cellular level occurs in the form of ions/molecules alone, the nanomaterial alone, or a combination of the two. For nanomaterials that dissolve, the released ions or molecules can be seen as breakdown products, providing options to justify read-across. Note that the fate/toxicokinetics should be taken into account as the ions/molecules may be formed at specific biological sites. However, dissolution rate in pure water does not necessarily correlate to dissolution rates in biological fluids or other aqueous media (Schoeters and Verougstraete 2007; Lowry et al., 2012). Therefore the strategy for information and data gathering should indicate the need to assess dissolution in biological fluids relevant to the organ at the point of exposure (e.g. sweat for skin (Di Cristo et al., 2022a), lung lining fluid for the lung (Murphy et al., 2021; Braakhuis et al., 2021) and a sequential combination of oro-gastro-intestinal fluids for the oral route of exposure (Di Cristo et al., 2021) or the relevant environmental compartment (e.g. aqueous (Cross et al., 2024), reconstituted water for sediment, (Little et al., 2021), Hoagland solution for plants (Song et al., 2022). Furthermore, for those particles which do not dissolve instantaneously in these different tissue fluids, it is also recommended that dissolution rate in phagolysosomal fluid is addressed. This allows to consider the release of metal ions within the acid pH of this intracellular compartment (Braakhuis et al., 2021; Di Cristo et al., 2021). This information helps to inform whether a bolus dose of the ions or molecules can be released inside the cell, avoiding the controlling uptake mechanisms of the cell membrane. When no or limited dissolution in lysosomal conditions occur, this is indicative of persistency in tissues and potential accumulation. Many nanomaterials consist of metals or metal oxides and

as such grouping according to their metal content may be relevant. Section 6.7.2 suggests the hypothesis *"properties are likely to be similar or follow a similar pattern if they have similar release rates or generate similar releases of the same metal moiety under comparable exposure circumstances"*. This hypothesis is relevant for nanomaterials for which the metal content can be dissolved. This is particularly true for metal or metal oxide nanomaterials that demonstrate dissolution within the human body or the environment within hours or days of exposure, allowing the metal ions to contribute to the mechanism of toxicity (Braakhuis et al., 2021; Di Cristo et al., 2021). As such dissolution rate is considered a key starting point in many nanomaterial grouping hypotheses (Murphy et al., 2023). For metal or metal oxide nanomaterials, the use of metal salts which allow exposure of cellular test systems to the same metallic ions could therefore be included as presumed group members for the similarity assessment of dissolution, fate, toxicokinetics and hazard.

- **Transformation:** Transformation at different life cycle stages can alter the physicochemical characteristics of nanomaterials and hence their fate, toxicokinetics and hazard (Williams et al., 2019; Spurgeon et al., 2020). The potential for transformation at relevant life cycle stages should be incorporated into the grouping hypothesis. Examples of transformations include dissolution, agglomeration, (bio)degradation or metabolism, or accumulation of a coating.
- **(Bio)degradation or metabolism:** Nanomaterials may be (bio)degraded or metabolised to substances with a known mode of action or AOP. Such degradation products or metabolites should be included within the grouping hypothesis and addressed as part of the data gathering and assessment process.

6.8.4. Limitations with grouping approaches for nanomaterials

Challenges with testing methods for characterisation and hazard assessment of nanomaterials

In the past, grouping of nanomaterials has been prevented by limitations in the availability of standardised methods for characterisation and hazard assessment (Sellers et al., 2015). However, this issue has improved over time via the activities of standardisation bodies (e.g. OECD, ISO, and national standardisation bodies) and internationally funded research projects. For example, OECD has generated test guidelines on the dispersion stability of manufactured nanomaterials (OECD, 2017), nanomaterials particle size (OECD, 2022), the determination of the volume specific surface area (OECD, 2022), and the determination of the hydrophobicity index (OECD, 2023).

For methods to assess physicochemical properties, fate, toxicokinetic and hazard there is sometimes uncertainty about their suitability due to variability of measurements (both for characterisation and hazard assessment), which means more work is either needed to improve and standardise the methods, or to develop alternatives (Basei et al., 2022). For such reasons there can be considerable variations between apparently comparable data from different sources (Comandella et al., 2020). Therefore, users need to assess the variability and completeness of their data to support a valid similarity assessment (see also (Marchese Robinson et al., 2016).

The nanomaterial-specific physicochemical properties also mean that conventional standardised hazard testing methods might not be appropriate (Sellers et al., 2015), either because they do not measure a hazard that is appropriate to the nanomaterial, because these methods have not been developed and adopted to meet the requirements for nanomaterials, or because the nanomaterial might interfere in the test method (Kroll et al., 2012). With respect to interference by nanomaterials in assays, this can be caused by a range of factors (Worth et al., 2017). For example, the large surface area of nanomaterials means

that they can adsorb biological molecules preventing their detection, or reagents preventing their participation in an assay reaction (Brown et al., 2010). The light absorbing, reflecting, or emitting properties of the particle might result in inaccurate detection of light signals associated with specific assays (Stone et al., 2009).

In addition, the way the particle is delivered into an experimental system can influence the data generated. The best protocol to disperse nanomaterials for hazard assessment remains an ongoing debate, which is influenced by factors such as the route of exposure, the environmental compartment, the dispersant used, agglomeration and the stability of the nanomaterial suspension generated, the use of sonication and the addition of organic matter (OECD, 2017f; OECD 2020m). It is worth noting that the updated "OECD Guidance on Sample Preparation and Dosimetry" was published in 2025 (OECD, 2025c).

The dispersion stability also influences the dose delivered to a particular test system (Cohen et al., 2013; DeLoid et al., 2014). For example, in an *in vitro* cell culture, particle size, density and agglomeration status of a suspended nanomaterial all influence their deposition rate and therefore the dose reaching the cells at the bottom of the culture dish. Conversely, for species such as *Daphnia magna* which live in the water column, settling will reduce the availability of the nanomaterials to the organisms and therefore influence the dose (Kennedy et al., 2017). The higher the nominal concentration, the higher the probability of an instable dispersion. Furthermore, in aquatic test systems, settling will result in a concentration gradient within the test column which will lead to altered exposure concentrations and unclear exposure situations based on the behaviour of the test organism (e.g. for pelagic organisms like daphnids or fish can move vertical through the test vessel). This makes increased frequency of monitoring and characterisation of exposure indispensable for reliability of toxicological data.

Some work has been done to assess the suitability of different assays for testing nanomaterials (Rasmussen et al., 2019); such methods can be more confidently used providing they have been assessed for nanomaterials relevant to the group under investigation. If any method has not been assessed for suitability using the nanomaterials being assessed for similarity, then either this assessment needs to be conducted, or the data may not be suitable for use in grouping.

Generally, data on nanomaterials need to be carefully evaluated and interpreted, taking into account the caveats described above, and keeping in mind that many methods used to generate the data are still under development and not standardised or validated yet.

Other issues related to grouping of nanomaterials

Physicochemical characteristics which exist as wide distributions (e.g. range of sizes, rates of dissolution) can be a problem for grouping. The distribution of physicochemical properties for nanomaterials means that, like UVCB categories, nanomaterial groups can be visualised more as a "cloud" or "sphere" of overlapping chemical and physical properties, rather than demonstrating a clear trend. Taking size as an example, it could be difficult to group all particle sizes together as fate and toxicokinetic is influenced by further characteristics as described above, instead individual size ranges may be required for different groups.

Understanding the impact of data adequacy/reliability and reproducibility is essential to assign a cut-off (threshold) value for a group or for the range of a floating band or for assessing whether materials can be considered similar (Basei et al., 2022; Comandella et al., 2020). Reproducibility checks on basic physicochemical information, demonstrate that these measurements can vary according to the methodology used (Comandella et al., 2020; Cross et al., 2021). This is important as it supports well-defined boundaries of the group, after the successful demonstration of similarity.

Similarity assessments often rely on properties that describe an interaction of the nanomaterial with its environment, such as reactivity (Seleci et al., 2022) or dissolution rates in specific media (Keller et al., 2021). These properties are typically classified as extrinsic, for which reproducibility information is often

currently limited to a range of 1.5-fold to 2-fold. While it is essential to consider these extrinsic properties for grouping hypotheses, it is equally important to seek better data with less variability. Available data from databases should ideally include this reproducibility information; if such data is lacking, its use may need to be reevaluated (Mancardi et al., 2023). This approach does not suggest ignoring extrinsic properties, as they remain vital for understanding nanomaterial behaviour, but rather emphasises the need for more reliable data to support robust grouping and read-across.

The suitability of multidimensional analysis of similarity of different nanomaterials across multiple descriptors is not yet sufficiently robust to support hazard assessment (Jeliazkova et al., 2022a; NIOSH, 2021). Such methods could be developed by using the expanding data sets that become available from larger scale research projects (such as collected in the Nanosafety Data Interface[134]). Such data will require analysis via the existing similarity approaches for comparison with the multidimensional approaches, as well as a clear, evidence-based understanding of the MOA driving the key hazard endpoint.

Grouping approaches so far are mostly applied to simple, single component nanomaterials, but they might not yet be suitable for more complex compositions as might be expected in more complex multicomponent nanomaterials or advanced materials (OECD, 2023c).

[134] Nanosafety Data Interface: https://enm-legacy.adma.ai/

6.8.5. Section 6.8 Annex – GRACIOUS Framework examples for generating a grouping hypothesis

- These case studies are illustrative examples and have not been reviewed by any OECD group.
- These examples should not be interpreted as reflecting official regulatory decisions made by OECD Member Countries.

A number of case studies using the GRACIOUS Framework (Stone et al., 2020) have been compiled as examples how to approach the grouping of nanomaterials, a few examples of which are provided below (Table 25). It is important to note that neither these examples nor the GRACIOUS Framework have been reviewed by OECD.

The case studies show that the GRACIOUS Framework can help the generation of a grouping hypothesis for nanomaterials, the strategic and logical data gathering needed to test the hypothesis, as well as using quantitative similarity methods in order to reduce reliance on expert judgement. A wide range of human health and environmental grouping hypotheses are proposed for a range of different nanomaterial types. For human health, the available case studies focus on apical hazards at the point of entry into the body following inhalation, ingestion, and dermal application. Further work is required to generate the data needed to apply grouping to distal effects. Furthermore, less evidence is available to support the generation of grouping hypotheses relevant to environmental hazard endpoints, while hypotheses are proposed for aquatic, sediment, and soil environments.

It should be noted that the case studies describe approaches to categorise nanomaterials in relation to their effects and MOA, not necessarily describing endpoint-specific grouping and read-across for data gap filling.

Table 25. Examples of case studies that have employed the GRACIOUS Framework by (Stone et al., 2020) to group nanomaterials for either human or environmental hazards

Publication	Relevant exposures	Materials	Comment
(Murphy et al., 2021)	Human Occupational Inhalation	Multiwalled carbon nanotubes (MWCNT)	Relevant to high aspect ratio (fiber like) nanomaterials with the potential to induce mesothelioma
(Di Cristo et al., 2022c)	Human Consumer or occupational Oral ingestion	Silica nanomaterials	Relevant to gradually dissolving nanomaterials that are ingested
(Song et al., 2022)	Environment Aquatic/soil	Metal or metal oxide nanomaterials	Assessing similarities between metals in relation to impacts on lettuce biomass and route elongation.
(Cross et al., 2024)	Environment Aquatic	Zinc oxide nanomaterials	Use of fate and ecotoxicity data to group nanomaterials in the aquatic environment

7 Reporting Formats for Analogue and Category Evaluations

Box 7.1. Chapter 7 summary

This chapter:

- Proposes revised reporting formats for analogue and chemical category approaches to ensure they adequately document and justify the read-across being performed.

  Although the original format described in the last two editions of this Guidance was informed by experiences gained from EU REACH, the IATA Case Studies[135] project has broadened the decision contexts and the types of data streams that form the basis for substantiating current read-across approaches. Accordingly, the reporting formats described in this chapter aim to consider where HTS/HCS and omics data as supporting information fit as well as ways in which read-across uncertainty can be characterised and documented. A critical aspect is that the reporting formats herein should be flexible in terms of how information can be summarised and documented. Further, the formats are intended to be modular to accommodate changes in the types of supporting information that may emerge in the future. Tables and reporting formats are provided in Annexes X-Y and are likely to be updated in the future.

The following sections detail the aspects to capture for the analogue and category approach separately, though it is recognised that there is a great deal of commonality in both. Although certain regulatory authorities may have specific information requirements, a core set of requirements is likely to be the same across many jurisdictions.

The reporting format follows a logical framework commensurate with the workflows described in Chapters 4 and 5. Every case should contain the following elements as appropriate:

- The **scope of application** of the analogue/category approach: the purpose for which the approach is applied as well as a list of endpoints where data gap filling is being proposed.
- The **analogue/category definition**: summary of common features; boundaries (e.g. number of carbon atoms); physicochemical properties, if applicable (e.g. boiling point); allowed variations in chemical structure; and if known, any restrictions (e.g. variations that would change the effects of a chemical significantly compared to the other chemicals in the case of a category approach). The analogue/category description will determine the applicability of the read-across hypothesis.

[135] http://www.oecd.org/chemicalsafety/risk-assessment/iata-integrated-approaches-to-testing-and-assessment.htm

- The **hypothesis** for identifying analogues/category members, where applicable: the structural, physicochemical, bioactivity, conventional toxicological and/or ADME/TK (including metabolic) similarities (e.g. functional group, moiety of concern, carbon chain length, common metabolic or degradation pathways and products, physicochemical properties such as boiling point or *log* K_{ow}, discussion of presence or absence of molecular and/or conventional toxicological effects).
- The **analogue/category approach justification**: all pre-existing experimental or other (e.g. literature) evidence that can support the hypothesis defined in the previous step. This could be similar effects in lower tier studies where these exist, evidence from computational and non-computational theoretic models, common bioavailability and reactivity profiles (empirical and/or *in silico*), similarity in bioactivity arising from *in vitro* or other high throughput screening studies or from conventional *in vivo* bridging studies, a common MOA or AOP. An endpoint-by-endpoint justification is required to demonstrate that a trend or similarity exists between the target/source analogue or members within the category. Demonstrating these trends is particularly relevant for the endpoint(s) where there are data gaps to be filled but this evaluation should not necessarily be restricted to just those endpoints.
- In the case of categories, the existence of subcategories and the **rationale for subcategorising** (e.g. existence of a threshold for certain physicochemical properties impacting solubility or bioavailability of category members and thus hazards for given endpoints).
- The **technique used to fill the data gap(s)**: indication for each endpoint and chemical of the availability of experimental data and for the target endpoints where no data were available, how was the data gap filling applied (e.g. read-across from most similar category member, worst case scenario, trend analysis, similarity weighted average), with justification for the strategy.

Generally, the aspects discussed in the other chapters of this guidance should be taken into consideration and reported adequately. To guide the user, the sections in the reporting format described below for the analogue and category approaches refer to the related steps of the stepwise approaches described in Chapters 4 and 5.

Reporting format for analogue approach

The following reporting format should not be viewed as a strict structure, rather it identifies the information that should be included in this type of report. Specific requirements related to the scope or specific regulatory requirements may apply and need to be taken into consideration.

1. Abstract/Synopsis/Executive summary

Provide a brief overview (for example 300 words) of the context for forming the analogue approach and its potential use including:

- Aim of the analogue approach and decision context
- Target chemical (identifiers/short description)
- Endpoints considered/route of exposure
- Source chemical(s) (identifiers/short description)
- Analogue definition and short statement on hypothesis/basis for read-across

2. Purpose

2.1. Problem formulation

Indicate the purpose of the grouping and read-across, the decision context and intended application. This relates to *Step 0* in the analogue workflow in Chapter 4 (see Section 4.2.1) which describes the problem formulation and decision context.

This may include:

a) Screening for priority setting for subsequent data collection efforts
b) Hazard identification
c) Risk assessment
d) Other or a combination of different purposes.

This will be pertinent when characterising and documenting the uncertainties later. The description of the purpose of use is important in considering whether the residual uncertainty of the analogue approach is acceptable.

2.2. Target substance

Provide the chemical descriptors and common identifiers, including CAS number, DSSTox substance identifier (DTXSID), EC number, name, composition, and structure as far as possible for the target substance (see Section **Error! Reference source not found.**). Structural information may be characterised in various ways, such as 2D chemical depictions, SMILES notation, and InChl code.

Provide purity/impurity profiles for the target substance, including their likely impact on the relevant endpoints. To the extent possible, discuss what influence these impurities are expected to have on physicochemical parameters, fate and (eco)toxicity, and in turn how these may impact the read-across being performed.

2.3. Endpoints considered/route of exposure

List the endpoints and route of exposure for which the analogue approach is being applied. This relates to Step 1 in the analogue workflow in Chapter 4 (see Section 4.2.2).

Depending on functional group(s), reactivity, bioactivity, metabolism, MOA, the analogue approach may apply for some endpoints only (e.g. acute effects only), in which case this should be specified and justified. Endpoint-specific considerations/approaches may be needed if more than one endpoint is addressed by the read-across.

3. Source substance(s)

This is the final result of Steps 3 "*Identify analogues"* (see Section 4.2.4) to 6 *"Evaluate the analogues"* (see Section 4.2.7), i.e. the source substance(s) to be used for the read-across.

For transparency, the methods and parameters used for analogue identification can be described in an Annex, if applicable. Analogue selection is performed after evaluation of the data and the adequacy of the analogues, based on an established rationale (hypothesis) that informs the applicability domain.

3.1 Source substance(s)

Provide the chemical descriptors and common identifiers for the source substance(s) as comprehensively as possible in the same manner as for the target substance. This information is critical to draw inferences between the target and the source substance(s).

Provide purity/impurity profiles for the source substance(s), including their likely impact on the relevant endpoints. To the extent possible, discuss what influence these impurities are expected to have on physicochemical parameters, fate and (eco)toxicology, and in turn how these may impact the read-across being performed.

3.2. Applicability domain

Provide a clear description of the applicability domain and structural boundaries. Describe the molecular structure a substance must have to be considered suitable as a source substance. A summary of common features, the boundaries (e.g. number of carbon atoms), physicochemical properties if applicable (e.g. boiling point), allowed variations in chemical structure, and if known, any restrictions (e.g. variations that would change the effects of a chemical significantly compared to the target chemical), common metabolic/degradation pathways etc. Clear inclusion and exclusion criteria should be described to define the applicability domain of the analogue approach.

The justification for the inclusion and/or exclusion rules should be discussed under 5 *"Justification of the analogue read-across approach"* below.

4. Summary of the available information

Summarise the data available (as collected or generated as supporting information in iteration of the workflow to support the hypothesis, if applicable). This relates to Steps 4 and 5 in the analogue workflow in Chapter 4 (see Sections 4.2.5 and 4.2.6). Briefly list the test methods or data sources used for gathering the data for the target and source substances.

4.1. Data matrix

Provide a matrix of data (endpoints versus target and source chemicals).

The target substance and the analogue(s) (source substance(s)) are listed on one axis as columns and the target endpoint(s) and all other endpoints and information on the other axis (as rows). If multiple analogues are identified, these could be arranged in a suitable order (e.g. according to MW or *log* K_{ow}).

Include physicochemical properties which are a critical determinant to the environmental and health properties of a substance affecting bioavailability, environmental fate, and thus the ecotoxicity and toxicity of a chemical (see Section **Error! Reference source not found.**).

All source substances relevant to the target substance should be included, as should all available data.

In the cells of the data matrix the result type should be indicated, for example:

- Experimental result
- Experimental study planned
- (Q)SAR/expert system prediction
- HTS/HCS or omics data
- AOP

If multiple results are available, the reliable key study results should be noted in the data matrix. Additional information could be provided in an Annex, with comments on the reliability.

Table 26 gives an indicative listing of the elements that should be captured. An example Excel template for the data matrix can also be found in the Appendix.

For some approaches, not all data will be available; and in other instances, the data available will be too complex to be inserted into a table.

Complex data, such as HTS/HCS or omics data may be better represented separately for example using a heatmap, stacked bar plot, radial line graph, or a multi-level donut chart to illustrate consistency of profiles between substances. Data could be summarised or aggregated by biological targets to facilitate ease of interpretation. For the particular case of omics data, the OECD Omics Reporting Framework (OORF) provides guidance (OECD 2023a). Further guidance on reporting is available in the Chemical Grouping Application Reporting Module (CG-ARM), see Appendix.

In the matrix, data gaps being filled by read-across (indicating the approach applied) from the source analogues should be highlighted.

If HTS/HCS or omics data are being used then indicate how the data is being used – to identify analogues, to justify bioactivity similarity, and/or to provide mechanistic information related to the endpoint being read across.

It is important to be transparent about selection of source substances and source studies in order to avoid bias (see Annexes I, III to the report).

The table in this report section should only show the final source substances selected for the read-across approach. Other identified possible analogue substances can be listed in an extended table in an Annex, including the assessment and reason for not being taken forward as source substances (see Step 6 of the workflow in Chapter 4, Section 4.2.7).

An assessment of the data quality is provided in the description of uncertainties below (Point 6.1).

Table 26. Example data matrix, analogue approach

Chemical ID				
		Target	Source1	(Source2, ...)
CAS				
Name				
Structure				
Summary of data gap filling				
		Target	Source1	(Source2, ...)
Target endpoint1	Experimental result		value, unit, test method (e.g. test guideline)	value, unit, test method (e.g. test guideline)
	Read-across result	derived result		
Target endpoint2	Experimental result		value, unit, test method (e.g. test guideline)	value, unit, test method (e.g. test guideline)
	Read-across result	derived result		
Physicochemical data				
Melting point				
Boiling point				
Density				
log K_{ow} (measured value)				
log K_{ow} (calculated value)				

…				
Kinetics*				
Absorption				
Distribution				
Metabolism				
Excretion				
Molecular profiling				
Parent chemical	Profiler 1 (name, version)			
	Expert system 1 (name, version)			
	…			
Metabolite**	Profiler 1 (name, version)			
	Expert system 1 (name, version)			
	...			
Supporting data related to the target endpoint(s)***				
		Target	Source1	(Source2, ...)
Endpoint 1	Method A			
	Method B			
	…			
	…			
	…			
Endpoint 2	Method C			
	Method D			
	…			
	…			
	…			
Endpoint 3	Method E			
	Method F			
	…			
	…			
	…			
Endpoint …	…			
	…			
	…			
Other data	…			
	…			
	…			

* General outline of relative comparative kinetics
** More relevant metabolite such as toxicant
*** *in vivo*, *in vitro*, *in chemico*, *in silico*, Defined Approach, Battery approach, other data, …

4.2. Description of information sources

In order to keep the core report shorter, details on the information sources and methods used to generate the data can be provided separately in an appendix to the report.

Briefly describe the methods/information sources used for the data matrix or provide reference to an existing description such as scientific literature, OECD Guideline for the Testing of Chemicals, or another guideline. If (Q)SAR data are included, provide name and version of model being used for deriving predictions.

For nanomaterials specifically, the description of methodologies for measurements of the relevant parameters is useful, including differences between the methodologies if several used for the same parameter, if applicable, as well as identification which parameters are relevant to which endpoints.

5. Justification of the analogue read-across approach

5.1. Analogue identification and selection

Relates to Step 6 – Adequacy of analogues of the workflow in Chapter 4, Section 4.2.7.

There are several different ways to identify potential analogues as source substances with data with which the target substance can be compared. It is usually an iterative process. Whereas it is well established now that structural similarity is only one criterion used to identify and evaluate the suitability of analogues for read-across, structural similarity can be a pragmatic first step in identifying promising analogues that could be expected to exhibit similarity in activity e.g. (eco)toxicity, environmental fate (see Step 3 of the workflow in Chapter 4, Section 4.2.4). Whilst considering similarity, it is pertinent to consider the impact of any dissimilarity in the approach (structure, bioactivity, etc.) and consider how these may affect the grouping and read-across strategy. Importantly, overall, a hypothesis should be established, why the target and source substances are sufficiently similar to conduct a read-across approach and use the data from the source to fill the data gap for the target (see Point 5.2 below).

Provide the hypothesis used to identify the source analogues, if applicable. Details on the strategy and tools used to identify possible analogues can be described in an annex.

This may include considerations on the structural, physicochemical, bioactivity, conventional toxicological and/or ADME/TK (including metabolic) similarities (e.g. functional group, moiety of concern, carbon chain length, common metabolic or degradation pathways and products, common physicochemical properties such as boiling point or *log* K_{ow}).

Also, in the case that the search strategy initially was not based on a specific rationale, describe the characteristics that a substance must have to be considered a suitable source analogue, i.e. considerations for analogue selection. A discussion of the presence or absence of molecular and/or conventional toxicological effects and considerations on the impact of structural differences is important (see Sections 4.2.4 and 4.2.7)

5.2. Analogue read-across approach

Provide the justification showing that the source substance is considered suitable and relevant in terms of the problem formulation, and the read-across approach is adequate. This relates to Step 7 of the workflow in Chapter 4, Section 4.2.8.

Aspects described in Chapter 3, Section **Error! Reference source not found.** should be taken into consideration.

Based on the available experimental data and other information, including basic physicochemical properties, available toxicity and ecotoxicity data from the source and the target substances, and

knowledge of metabolism or MOA or AOP, summarise how these results substantiate the hypothesis and read-across proposed. It should be explained why the read-across between the source and target substance is justified, and why the analogue(s) used are adequate and provide sufficiently robust information to characterise the hazard endpoint(s) considered.

Information from positive and negative control chemicals (sometimes referred to as anchor chemicals) could also be considered here. The data should also show that dissimilarities in chemical structure between sources and target, e.g. functional groups not common to source and target substances, do not affect the anticipated toxicity (see also Step 3 of the workflow in Chapter 4, Section 4.2.4). The available information in the data matrix reported under Point 4.1 above should support the justification for the read-across. This could include evidence from computational and non-computational theoretic models, common bioavailability and reactivity profiles (empirical and/or *in silico*), similarity in bioactivity arising from *in vitro* or other high throughput screening or omics studies or from conventional *in vivo* bridging studies, a common MOA or AOP.

For the particular case of omics data, the OECD Chemical Grouping Application Reporting Module (CG-ARM), see Appendix, describes how to report a well-structured narrative for grouping chemicals using this data type, and cross-references to other reporting modules within the OECD Omics Reporting Framework (OECD 2023a). It comprises a rationale for the experimental design and use of omics data for grouping, a description of the omics data and metadata, the methods used to assess the bioactivity similarity of chemicals, the results of the similarity assessment to support the analogue justification, and a description of the uncertainties associated with the omics data and results.

Provide the justification on an endpoint basis. An endpoint-by-endpoint justification is required to demonstrate that a similarity exists between the target and source analogue and read-across can be performed. This should not be restricted to only those endpoints where read-across is being proposed. Provide the hypothesis for why the endpoint specific read-across can be performed with reference to the similarity contexts for the source analogue(s). This may include a discussion for the structural, reactivity, metabolic, physicochemical, bioactivity, mechanistic, and toxicological similarities between the target chemical and the analogue(s) (if not discussed before). Respective evidence with the relevant supporting information should be provided (the relevance of the information for each endpoint should be explained).

If there is a mechanistic reasoning to the read-across, describe the foreseen MOA or AOP for source and target chemicals and if relevant describe the influence of the mode of administration of the source chemical (oral, dermal, inhalation) and its relevance in relation to the physical form of the target chemical. See Sections 2.4 and 2.5 for more guidance.

The graphical representation of the AOP if relevant would also be helpful, as well as key references.

For nanomaterials

The parameters to be considered for read-across of nanomaterials and their relevance for human health and environmental endpoints are for example surface chemistry, size, shape, and surface area, along with physicochemical properties.

An explanation of which parameters are critical for the analogue approach hypothesis should be provided. Properties and characteristics of nanomaterials relevant for hypothesis generation are described in Section 6.8.3 (though the list is not exhaustive).[136]

(Homo- and hetero-) Aggregation and agglomeration behaviour would also be critical to evaluate bioavailability. Surface chemistry is essential, but core material chemical identity is also of high importance.

See Section 6.8 *"Grouping of nanomaterials"*.

More information is also available at (ECHA, 2019).

6. Read-across conclusion

Relates to Step 8 of the workflow in Chapter 4, Section 4.2.9, and Step 9, Section 4.2.10.

6.1 Description of uncertainties

Describe the uncertainties of the different aspects for the grouping and read- across. For the given purpose, the consideration of uncertainty may start from the choice of hypothesis. Aspects can include uncertainty and confidence associated with all types of data used for supporting the read-across hypothesis, the underlying data used to be read across from the source chemicals, e.g. type, quality, and relevance (see Step 6 - Availability and quality of data of the workflow in Chapter 4, Section 4.2.7), as well as assumptions used to develop the similarity rationale of the analogue substances, the applicability domain. Uncertainty aspects are described in Chapter 2, Section 2.5, main aspects to consider for uncertainty assessment are summarised in Figure 3.

Table 27 provides an example of reporting uncertainties related to the aspect of the similarity rationale. It can be modified as appropriate. It is also recommended to state what is not addressed.

Examples of templates/systematic approaches to assess uncertainties are also described in Section 2.6.2.

[136] For parameters to be considered for grouping of nanomaterials see also (ECHA, 2019) Appendix I, Table 2 "*Key physico-chemical parameters to be considered for grouping and read-across of nanoforms and their relevance for human health and environmental endpoints".*

Table 27. Example of uncertainty reporting related to the similarity rationale

Similarity Rationale	Data quality (low, medium, high)	Strength of evidence (low, medium, high)	Comments
Structural similarity			
Physicochemical property similarity			
ADME/TK including metabolic similarity			
Chemical reactivity profile			
Bioactivity similarity			
Conventional toxicological profile			
MOA similarity			
Overall uncertainty			Concordance and WoE of all data used for justifying the hypothesis

For nanomaterials

In addition to the above-mentioned aspects, the following should be considered in the characterisation of uncertainties related to the analogue approach for nanomaterials:

- Complexity of nanostructures similarity
- Identity characterisation of the nanomaterials
- Variability of the measurements, test system relevance for nanomaterials and possible nanospecific artefacts in assays

See Section 6.8 *"Grouping of nanomaterials"*.

More information is also available at (ECHA, 2019).

6.2. Result of read-across and data gap filling

Describe the result of the read-across approach and the data gap filling applied. Reflect on the result in view of the scope, addressing the problem formulation, rationale, and whether the uncertainty is acceptable.

The conclusion of the analogue approach should be documented for the individual endpoints (data gaps).

- Provide the strategy and rationale used to fill the data gap and integrated conclusion of data gap filling, including description how the data gap is filled (e.g. average, most sensitive, similarity weighted, qualitative).
- Discuss the remaining uncertainties (see Chapter 2), in view of the problem formulation and how they might be addressed/overcome or whether the confidence in the result is sufficient for the purpose.
- Finally, provide a short conclusion summarising the outcome of the evaluation and the final result.

Detailed information on the data used for data gap filling (result, test type, study design) should be provided.

7. Annexes to the reported analogue approach

Inclusion of this additional information depends on the context. Specific requirements related to the scope or specific regulatory requirements may apply and need to be taken into consideration.

- **Annex I: Methods and search strategies used to identify analogue substances**

  This can be added for transparency, in order to understand how analogues were identified.

  State for example computational tools and parameters used for the search. It should be noted that the identification strategy can be an exploratory and iterative process, and analogue selection also depends on the availability of data. Analogue selection is performed after evaluation of the data and the adequacy of the analogues, based on an established rationale (hypothesis) that informs the applicability domain.

  For transparency and showing no bias in the analogue selection, excluded potential analogues should be mentioned together with the reason why they were excluded. This can also include lack of data available for these analogues.

- **Annex II: Information sources and methods of data generation**

  Describe the experimental or computational methods with which the data available for the substances were generated. Note that this does not necessarily imply that data were generated by the user applying the read-across. Method descriptions allow to understand the methods and parameters applied and are also important for an assessment of the data quality and relevance for the endpoint(s) and purpose considered.

  Specifically for nanomaterials, if there are several methods used for measuring specific parameters, it is helpful to note what the differences between the methodologies are in terms of determining these parameters (e.g. leading to different levels of uncertainty), if applicable, for the comparison of analogue substance properties and assessment of similarity.

- **Annex III: Assessment of the available data**

  Provide detailed discussion of available test results for individual endpoints (i.e. discussion of the selection of key studies, reliability of the experimental data, variability of experimental results between source and target substances, the quality of the data estimated by external computational approaches etc.).

  Additional (supporting) data available for the substances can be provided, which are not listed in the data matrix, together with an assessment of the data quality and relevance. For transparency and showing no bias in the data selection, possible excluded studies should be mentioned together with the reason why they were excluded.

Reporting format for chemical categories

The following reporting format should not be viewed as a strict structure, rather it identifies the information that should be included in this type of report. Specific requirements related to the scope or specific regulatory requirements may apply and need to be taken into consideration.

1. Abstract/Synopsis/Executive summary

Provide a brief overview (for example 300 words) of the context for forming the category approach and its potential use including:

- Aim of the category approach and decision context
- Target chemical (identifiers/short description)
- Endpoints considered/route of exposure
- Category members (identifiers/short description)
- Category definition and short statement on hypothesis/basis for read-across

2. Purpose

2.1. Problem formulation

Indicate the purpose of the grouping and read-across, the decision context and intended application. This relates to Step 0 of the workflow in Chapter 5, Section 5.2.1 which describes the problem formulation and decision context.

This may include:

a. Screening for priority setting for subsequent data collection efforts
b. Hazard identification
c. Risk assessment
d. Other or a combination of different purposes

This will be pertinent when characterising and documenting the uncertainties later. The description of the purpose of use is important in considering whether the residual uncertainty of the category approach is acceptable.

2.2. Target substance

Provide the chemical descriptors and common identifiers, including CAS number, DSSTox substance identifier (DTXSID), EC number, name, composition, and structure as far as possible for the target substance (see Section **Error! Reference source not found.**). Structural information may be characterised in various ways, such as 2D chemical depictions, SMILES notation, and InChl code.

Provide purity/impurity profiles for the target substance, including their likely impact on the relevant endpoints. To the extent possible, discuss what influence these impurities are expected to have on physicochemical parameters, fate and (eco)toxicity, and in turn how these may impact the read-across being performed.

2.3. Endpoints considered/route of exposure

List the endpoints and route of exposure for which the category approach is being applied. This relates to Step 1 of the workflow in Chapter 5, Section 5.2.2.

Depending on functional group(s), reactivity, bioactivity, metabolism, MOA, the category approach may apply for some endpoints only (e.g. acute effects only), in which case this should be specified and justified. Endpoint-specific considerations/approaches may be needed if more than one endpoint is addressed by the read-across.

In addition, indicate if, for some endpoints, the category approach can only be applied to a subset of the members of the category (subcategories).

3. Category members

This is the final result of Steps 3 of workflow in Chapter 5, Section 5.2.4, to 6 ,Section 5.2.8 i.e. the category developed to be used for the read-across.

For transparency, the methods and parameters used for the development of the category can be described in an annex, if applicable.

3.1. Category members

For each category member, provide the chemical descriptors and common identifiers as comprehensively as possible in the same manner as for the target substance. Section 2.3.2 gives information on category membership and Section **Error! Reference source not found.** on chemical identity and compositional aspects.

As described in Section **Error! Reference source not found.**, impurities are important to consider and may relate to the source and manufacturing route of the substance under consideration. Consequently, it is of value to provide purity/impurity profiles for each member of the category, including their likely impact on the category endpoints. To the extent possible, discuss what influence these impurities are expected to have on physicochemical parameters, fate and (eco)toxicity properties, and in turn how these may impact the read-across being performed. Considerations for UVCBs are provided in Chapter 6.6.

3.1.1. Compositional Information

As described in Section **Error! Reference source not found.**, while structure is critical for some categories, composition within the category members may be the critical factor for others. This is especially true for multi-constituent substance and UVCB categories and examples are given of categories based upon component analyses (see Table 5).

An approach is described for reporting upon and characterising category composition in Section 3.2.3 and example table for reporting upon category composition is also given (Table 6). Such a table should be considered to give the maximum clarity to the bounds of the category and the proposed read-across.

3.1.2. Physicochemical properties

As described in Section 3.2.3, physicochemical properties are a critical determinant to the environmental and health properties of a substance affecting bioavailability, environmental fate, and thus the toxicity and ecotoxicity of a chemical. Consequently, physicochemical properties across a category should be elaborated as part of its basic properties and reported in a table. A plot of a trend is usually very helpful to give clarity.

3.1.3. Example for structural matrix

The structural and functional elements and the relationship between the various category members need to be stated in a clear and unambiguous manner. Table 28 gives examples of how to build a matrix of the

category members and to present the elements that change or stay constant within the category, and to provide the most representative structures.

Table 28. Examples for structural matrix

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

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

- Salts
- Carbon number of chain
- Degree and nature of branching or occurrence of double bonds, functional groups, aromatics, cycles, heterocycles
- Moiety
- Valency
- Positioning of the common functional element

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

3.2. Applicability domain

Provide a clear description of the definition of the category, i.e. the applicability domain and structural boundaries. Describe the molecular structure a substance must have to be included in the category.

A summary of the common features of the category members, the boundaries (e.g. number of carbon atoms), physicochemical properties if applicable (e.g. boiling point), allowed variations in chemical structure, and if known, any restrictions (e.g. variations that would change the effects of a substance significantly compared to the other substances in the category) (see Section **Error! Reference source not found.**).

Identify the endpoint(s) for which the category approach is applied. Endpoint-specific considerations/approaches may be needed if more than one endpoint is addressed by the read-across.

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

Clearly indicate the boundaries of the category and for which substances the category does not hold, i.e. any restrictions such as variations that would change the effects of a chemical significantly compared to

the other chemicals of the category. The justification for the inclusion and/or exclusion rules should be discussed under Point 5 below.

4. Summary of the available information

Summarise the data available (as collected or generated as supporting information in iteration of the workflow to support the hypothesis, if applicable). This relates to Steps 4 and 5 of the category workflow in Chapter 5 (see Sections 5.2.5 and 5.2.6). Briefly list the test methods or data sources used for gathering the data for the target and category members.

4.1. Data matrix

Provide a matrix of data (endpoints versus target and category members).

The target substance and the category members are listed on one axis as columns and the target endpoint(s) and all other endpoints and information on the other axis (as rows). The matrix should be constructed with the category members arranged in a suitable order (e.g. according to molecular weight) to add clarity to trends in any relevant properties. For example, the ordering of the members should reflect a trend or progression within the category. In case of subcategories, these should be easily identifiable in the data matrix table.

Include physicochemical properties which are a critical determinant to the environmental and health properties of a substance affecting bioavailability, environmental fate, and thus the ecotoxicity and toxicity of a chemical (see Section **Error! Reference source not found.**).

All category members relevant to the target substance should be included, as should all available data. In the cells of the data matrix the result type should be indicated, for example:

- Experimental result
- Experimental study planned
- Trend analysis137
- (Q)SAR/expert system prediction
- HTS/HCS or omics data
- AOP

If multiple results are available, the reliable key study results should be noted in the data matrix. Additional information could be provided in an annex, with comments on the reliability.

Table 29 gives an indicative listing of the elements that should be captured. An example Excel template for the data matrix can also be found in the Appendix.

For some approaches, not all data will be available; and in other instances, the data available will be too complex to be inserted into a table.

Complex data, such as HTS/HCS or omics data may be better represented separately for example using a heatmap, stacked bar plot, radial line graph, or a multi-level donut chart to illustrate consistency of profiles between substances. Data could be summarised or aggregated by biological targets to facilitate ease of interpretation. For the particular case of omics data, the OECD Omics Reporting Framework (OORF)

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

provides guidance (OECD 2023a). Further guidance on reporting is available in the Chemical Grouping Application Reporting Module (CG-ARM), see Appendix.

In the matrix, data gaps filled by read-across from category members should be highlighted (indicating the approach applied, e.g. closest category members, average from category members, worst-case scenario). It is useful to indicate which category members are used for the read-across for each data gap.

If HTS/HCS or omics data are being used then indicate how the data is being used – to identify category members, to justify bioactivity similarity, and/or to provide mechanistic information related to the endpoint being read across.

It is important to be transparent about selection of category members and source studies in order to avoid bias (see Annexes I, III to the report).

The table in this report section should only show the final category members selected for the read-across approach. Other identified possible category members can be listed in an extended table in an Annex, including the assessment and reason for not being taken forward as source substances (see Step 6 of the workflow in Chapter 5, Section 5.2.7).

An assessment of the data quality is provided in the description of uncertainties below (Point 6.1).

Table 29. Example data matrix, chemical category

Chemical ID						
		Member 1	Member 2	Member 3	…	Member n
CAS						
Name						
Structure						
Summary of data gap filling						
		Member 1	Member 2	Member 3	…	Member n
Target endpoint1	Experimental result	value, unit, test method (e.g. test guideline)		value, unit, test method (e.g. test guideline)	…	value, unit, test method (e.g. test guideline)
	Read-across result		derived result		…	
Target endpoint2	Experimental result	value, unit, test method (e.g. test guideline)		value, unit, test method (e.g. test guideline)	…	value, unit, test method (e.g. test guideline)
	Read-across result		derived result		…	
Physicochemical data						
Melting point						
Boiling point						
Density						
log K_{ow} (measured value)						
log K_{ow} (calculated value)						
…						
Kinetics*						
Absorption						
Distribution						
Metabolism						
Excretion						
Molecular profiling						
Parent chemical	Profiler 1 (name, version)					
	Expert system 1 (name, version)					
	…					
Metabolite**	Profiler 1 (name, version)					
	Expert system 1 (name, version)					
	…					
Supporting data related to the target endpoint(s)***						
		Member 1	Member 2	Member 3	…	Member n
Endpoint 1	Method A					
	Method B					
	…					
	…					
	…					

Endpoint 2	Method C					
	Method D					
	…					
	…					
	…					
Endpoint 3	Method E					
	Method F					
	…					
	…					
	…					
Endpoint …	…					
	…					
	…					
Other data	…					
	…					
	…					

* General outline of relative comparative kinetics
** More relevant metabolite such as toxicant
*** *in vivo*, *in vitro*, *in chemico, in silico*, Defined Approach, Battery approach, other data, ...

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

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

4.2. Description of information sources

In order to keep the core report shorter, details on the information sources and methods used to generate the data can be provided separately in an appendix to the report.

Briefly describe the methods/information sources used for the data matrix or provide reference to an existing description such as scientific literature, OECD Guideline for the Testing of Chemicals, or another guideline. If (Q)SAR data are included, provide name and version of model being used for deriving predictions.

For nanomaterials specifically, the description of methodologies for measurements of the relevant parameters is useful, including differences between the methodologies if several used for the same parameter, if applicable, as well as identification which parameters are relevant to which endpoints.

5. Justification of the category read-across approach

5.1. Identification of category members and category development

Relates to Step 3 of the workflow in Chapter 5, Section 5.2.4, and Step 6 - Adequacy of category members, Section 5.2.7.

Provide the hypothesis used to identify the category members. Details on the tools used to identify possible category members can be described in an annex.

This may include considerations on the structural, physicochemical, bioactivity, conventional toxicological and/or ADME/TK (including metabolic) similarities (e.g. functional group, moiety of concern, carbon chain length, common metabolic or degradation pathways and products, common physicochemical properties such as boiling point or *log* K_{ow}).

Describe the characteristics that a substance must have to be considered a suitable category member, i.e. considerations for category member selection (see Section 5.2.4*).* A discussion of the presence or absence of molecular and/or conventional toxicological effects and considerations on the impact of structural differences is important.

Furthermore, an indication of the trend(s) for each endpoint should be included, if such trend exists, and what explains the trend observed (e.g. incremental structural changes) and whether such trend applies to the whole category or whether breakpoints/thresholds are to be expected (see 5.2.85.2.7). The existence of subcategories and the rationale for subcategorising (e.g. existence of thresholds in physicochemical properties impacting solubility or bioavailability of category members and thus hazards) should be provided and justified.

5.2. Category read-across approach

Provide the justification showing that the category members/the formed category are considered suitable and relevant in terms of the problem formulation, and the read-across approach is adequate. This relates to Step 7 of the workflow in Chapter 5, Section 5.2.8.

Aspects described in Chapter 3, Section **Error! Reference source not found.** should be taken into consideration.

Provide the hypothesis for why the category was formed: the hypothetical relational features of the category i.e. the chemical similarities (analogies), purported mechanisms and trends in properties and/or activities that are thought to collectively generate an association between the members (see Section 5.2.4). If there is a mechanistic reasoning to the category, describe the foreseen mode of action for each category member and if relevant describe the influence of the mode of administration, i.e. oral, dermal, inhalation (see for more information Section 2.4).

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

Based on the available experimental data and other information, including basic physicochemical properties, additional test results and molecular descriptor or profiler values that might have been generated for the assessment of this category, available toxicity and ecotoxicity data from category members and the target substance, and knowledge of metabolism or mode and/or mechanism of action or adverse outcome pathway, summarise how these results for each included substance in the category verify that the category is robust. It should be explained why the read-across between the category members and target substance is justified, and why the category members used are adequate and provide sufficiently robust information to characterise the hazard endpoint(s) considered.

This applies to the whole category, when no trend appears clearly across the category, or subcategories, if applicable. The strategy applied for the read-across should be stated and justified (e.g. read-across from the closest analogue, average from several analogues, or worst-case scenario).

Information from positive and negative control chemicals (sometimes referred to as anchor chemicals) could also be considered here. The data should also show that dissimilarities in chemical structure between category members and target, e.g. functional groups not common to all the (sub)category members do not affect the anticipated toxicity (see also Section 5.2.4). The available information in the data matrix reported

under Point 4.1. above should support the justification for the category and the read-across. This could include evidence from computational and non-computational theoretic models, common bioavailability and reactivity profiles (empirical and/or *in silico*), similarity in bioactivity arising from *in vitro* or other HTS/HCS or omics studies or from conventional *in vivo* bridging studies, a common MOA or AOP.

For the particular case of omics data, the OECD Chemical Grouping Application Reporting Module (CG-ARM), see Appendix, describes how to report a well-structured narrative for grouping chemicals using this data type, and cross references to other reporting modules within the OECD Omics Reporting Framework (OECD 2023a). It comprises a rationale for the experimental design and use of omics data for grouping, a description of the omics data and metadata, the methods used to assess the bioactivity similarity of chemicals, the results of the similarity assessment to support the category justification, and a description of the uncertainties associated with the omics data and results.

Provide the justification on an endpoint basis. An endpoint-by-endpoint justification is required to demonstrate that a similarity exists between the category members and read-across can be performed. This should not be restricted to only those endpoints where read-across is being proposed. Provide the hypothesis for why the endpoint specific read-across can be performed with reference to the similarity contexts for the category members. This may include a discussion for the structural, reactivity, metabolic, physicochemical, bioactivity, mechanistic and toxicological similarities between the target chemical and the analogue(s) (if not discussed before). Respective evidence with the relevant supporting information should be provided (the relevance of the information for each endpoint should be explained).

If there is a mechanistic reasoning to the read-across, describe the foreseen MOA or AOP for source and target chemicals and if relevant describe the influence of the mode of administration of the source chemical (oral, dermal, inhalation) and its relevance in relation to the physical form of the target chemical. See Sections 2.4 and 2.5 for more guidance.

The graphical representation of the AOP if relevant would also be helpful, as well as key references.

For nanomaterials

The parameters to be considered for read-across of nanomaterials and their relevance for human health and environmental endpoints are for example surface chemistry, size, shape, and surface area, along with physicochemical properties.

An explanation of which parameters are critical for the category approach hypothesis should be provided. Properties and characteristics of nanomaterials relevant for hypothesis generation are described in Section 6.8.3 (though the list is not exhaustive).[138]

(Homo- and hetero-) Aggregation and agglomeration behaviour would also be critical to evaluate bioavailability. Surface chemistry is essential, but core material chemical identity is also of high importance.

See Section 6.8 *"Grouping of nanomaterials"*.

More information is also available at (ECHA, 2019).

[138] For parameters to be considered for grouping of nanomaterials see also (ECHA, 2019) Appendix I, Table 2 "Key physicochemical parameters to be considered for grouping and read-across of nanoforms and their relevance for human health and environmental endpoints".

6. Read-across conclusion

Relates to Step 8 of the workflow in Chapter 5, Section 5.2.9, and Step 9, Section 5.2.10.

6.1. Description of uncertainties

Describe the uncertainties of the different aspects for the grouping and read- across. For the given purpose, the consideration of uncertainty may start from the choice of hypothesis. Aspects can include uncertainty and confidence associated with all type of data used for supporting the read-across hypothesis, the underlying data used to be read across from the source chemicals e.g. type, quality, and relevance (see Step 6 - Availability and quality of data of the workflow in Chapter 5 (Section 5.2.7), as well as assumptions used to develop the similarity rationale of the category members, the applicability domain. Uncertainty aspects are described in Chapter 2, Section 2.5, main aspects to consider for uncertainty assessment are summarised in Figure 3.

The following Table 30 provides an example of reporting uncertainties related to the aspect of the similarity rationale. It can be modified as appropriate. It is also recommended to state what is not addressed.

Examples of templates/systematic approaches to assess uncertainties are also described in Section 2.6.2.

Table 30. Example of uncertainty reporting related to the similarity rationale

Similarity Rationale	Data quality (low, medium, high)	Strength of evidence (low, medium, high)	Comments
Structural similarity			
Physicochemical property similarity			
ADME/TK including metabolic similarity			
Chemical reactivity profile			
Bioactivity similarity			
Conventional toxicological profile			
MOA similarity			
Overall uncertainty			Concordance and WoE of all data used for justifying the hypothesis

For nanomaterials

In addition to the above-mentioned aspects, the following should be considered in the characterisation of uncertainties related to the category approach for nanomaterials:

- Complexity of nanostructures similarity
- Identity characterisation of the nanomaterials
- Variability of the measurements, test system relevance for nanomaterials and possible nanospecific artefacts in assays

See Section 6.8 *"Grouping of nanomaterials"*.

More information is also available at (ECHA, 2019).

6.2. Result of read-across and data gap filling

Describe the result of the read-across approach and the data gap filling applied. Reflect on the result in view of the scope, addressing the problem formulation, rationale and whether the uncertainty is acceptable. If subcategories have been developed or if the category hypothesis specifies which endpoints are in scope, any data gap filling should be performed for the relevant category members only.

The conclusion of the category approach should be documented for the individual endpoints (data gaps).

- Provide the strategy and rationale used to fill the data gap and integrated conclusion of data gap filling, including description how the data gap is actually filled (e.g. average, most sensitive, similarity weighted, qualitative).
- Discuss the remaining uncertainties (see Chapter 2), in view of the problem formulation and how they might be addressed/overcome or whether the confidence in the result is sufficient for the purpose.
- Finally, provide a short conclusion summarising the outcome of the evaluation and the final result.

Detailed information on the data used for data gap filling (result, test type, study design) should be provided.

7. Annexes to the reported category approach

Inclusion of this additional information depends on the context. Specific requirements related to the scope or specific regulatory requirements may apply and need to be taken into consideration.

- **Annex I: Methods and strategies used to develop the category**

  This can be added for transparency, in order to understand how the category members were identified.

  State for example computational tools and parameters used for the search. It should be noted that the identification strategy can be an exploratory and iterative process, and selection of category members also depends on the availability of data. The category is formed after evaluation of the data and the adequacy of the (potential) category members, based on the established rationale (hypothesis) that informs the applicability domain.

  For transparency and showing no bias in the category member selection, excluded potential category members should be mentioned together with the reason why they were excluded. This

can also include lack of data available for these category members. Special consideration should be given to presumed outliers, if not already discussed in the context of subcategorisation.

- **Annex II: Information sources and methods of data generation**

  Describe the experimental or computational methods with which the data available for the substances were generated. Note that this does not necessarily imply that data were generated by the user applying the read-across. Method descriptions allow to understand the methods and parameters applied and are also important for an assessment of the data quality and relevance for the endpoint(s) and purpose considered.

  Specifically for nanomaterials, if there are several methods used for measuring specific parameters, it is helpful to note what the differences between the methodologies are in terms of determining these parameters (e.g. leading to different levels of uncertainty), if applicable, for the comparison of category member properties and assessment of similarity.

- **Annex III: Assessment of the available data**

  Provide detailed discussion of available test results for individual endpoints (i.e. discussion of the selection of key studies, reliability of the experimental data, variability of experimental results between category members and target substances, the quality of the data estimated by external computational approaches etc.).

  Additional (supporting) data available for the substances can be provided, which are not listed in the data matrix, together with an assessment of the data quality and relevance. For transparency and showing no bias in the data selection, possible excluded studies should be mentioned together with the reason why they were excluded.

References

Adams, M. et al. (2023), “Development of a CSRML version of the Analog identification Methodology (AIM) fragments and their evaluation within the Generalised Read-Across (GenRA) approach”, *Computational Toxicology*, Vol. 25, p. 100256, https://doi.org/10.1016/j.comtox.2022.100256.

Adams, W. et al. (2020), “Bioavailability Assessment of Metals in Freshwater Environments: A Historical Review”, *Environmental Toxicology and Chemistry*, Vol. 39/1, pp. 48-59, https://doi.org/10.1002/etc.4558.

Api, A. et al. (2022), “The RIFM approach to evaluating Natural Complex Substances (NCS)”, *Food and Chemical Toxicology*, Vol. 159, p. 112715, https://doi.org/10.1016/j.fct.2021.112715.

Aptula, A. et al. (2006), “Non-enzymatic glutathione reactivity and *in vitro* toxicity: A non-animal approach to skin sensitization”, *Toxicology in Vitro*, Vol. 20/2, pp. 239-247, https://doi.org/10.1016/j.tiv.2005.07.003.

Arts, J. et al. (2015), “A decision-making framework for the grouping and testing of nanomaterials (DF4nanoGrouping)”, *Regulatory Toxicology and Pharmacology*, Vol. 71/2, pp. S1-S27, https://doi.org/10.1016/j.yrtph.2015.03.007.

Arts, J. et al. (2016), “Case studies putting the decision-making framework for the grouping and testing of nanomaterials (DF4nanoGrouping) into practice”, *Regulatory Toxicology and Pharmacology*, Vol. 76, pp. 234-261, https://doi.org/10.1016/j.yrtph.2015.11.020.

Aschberger, K. et al. (2019), “Grouping of multi-walled carbon nanotubes to read-across genotoxicity: A case study to evaluate the applicability of regulatory guidance”, *Computational Toxicology*, Vol. 9, pp. 22-35, https://doi.org/10.1016/j.comtox.2018.10.001.

Attarilar, S. et al. (2020), “The Toxicity Phenomenon and the Related Occurrence in Metal and Metal Oxide Nanoparticles: A Brief Review From the Biomedical Perspective”, *Frontiers in Bioengineering and Biotechnology*, Vol. 8, https://doi.org/10.3389/fbioe.2020.00822.

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

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

ATSDR (2022), Guidance for Calculating Benzo(a)pyrene Equivalents for Cancer Evaluations of Polycyclic Aromatic Hydrocarbons.Atlanta, GA: U.S. Department of Health and Human Services, Public Health Service. April 14.

Ball, N. et al. (2014), “The challenge of using read-across within the EU REACH regulatory framework; how much uncertainty is too much? Dipropylene glycol methyl ether acetate, an exemplary case study”, *Regulatory Toxicology and Pharmacology*, Vol. 68/2, pp. 212-221, https://doi.org/10.1016/j.yrtph.2013.12.007.

Bal-Price, A. (2018), “Recommendation on test readiness criteria for new approach methods in toxicology: Exemplified for developmental neurotoxicity”, *ALTEX*, pp. 306-352,

https://doi.org/10.14573/altex.1712081.

Banerjee, A. et al. (2024), "Molecular similarity in chemical informatics and predictive toxicity modeling: from quantitative read-across (q-RA) to quantitative read-across structure–activity relationship (q-RASAR) with the application of machine learning", *Critical Reviews in Toxicology*, Vol. 54/9, pp. 659-684, https://doi.org/10.1080/10408444.2024.2386260.

Basei, G. et al. (2019), "Making use of available and emerging data to predict the hazards of engineered nanomaterials by means of *in silico* tools: A critical review", *NanoImpact*, Vol. 13, pp. 76-99, https://doi.org/10.1016/j.impact.2019.01.003.

Basei, G. et al. (2022), "A methodology for the automatic evaluation of data quality and completeness of nanomaterials for risk assessment purposes", *Nanotoxicology*, Vol. 16/2, pp. 195-216, https://doi.org/10.1080/17435390.2022.2065222.

Basei, G. et al. (2021), "A Weight of Evidence approach to classify nanomaterials according to the EU Classification, Labelling and Packaging Regulation criteria", *NanoImpact*, Vol. 24, p. 100359, https://doi.org/10.1016/j.impact.2021.100359.

BAuA (2013), German Federal Institute for Occupational Safety and Health. Announcement regarding hazardous substances. Manufactured nanomaterials: BekGS 527. BMBl 25:498–511, http://www.baua.de/de/Themen-von-A-Z/Gefahrstoffe/TRGS/Bekanntmachung-527.html

Beischlag, T. et al. (2008), "The Aryl Hydrocarbon Receptor Complex and the Control of Gene Expression", *Critical Reviews™ in Eukaryotic Gene Expression*, Vol. 18/3, pp. 207-250, https://doi.org/10.1615/critreveukargeneexpr.v18.i3.20.

Behl, M. and M. J. Hooth (2014), NTP Technical Report on the Toxicology Studies of Cobalt Metal (CASRN 7440-48-4) in F344/N Rats and B6C3F1/N Mice and Toxicology and Carcinogenesis Studies of Cobalt Metal in F344/NTac Rats and B6C3F1/N Mice (Inhalation Studies), NIEHS, https://doi.org/10.22427/ntp-tr-581.

Benfenati, E. et al. (2019), "Integrating *in silico* models and read-across methods for predicting toxicity of chemicals: A step-wise strategy", *Environment International*, Vol. 131, p. 105060, https://doi.org/10.1016/j.envint.2019.105060.

Benigni, R. et al. (2005), "Simple and α,β-unsaturated aldehydes: Correct prediction of genotoxic activity through structure-activity relationship models", *Environmental and Molecular Mutagenesis*, Vol. 46/4, pp. 268-280, https://doi.org/10.1002/em.20158.

Besser, J. et al. (2013), "Chronic toxicity of nickel-spiked freshwater sediments: Variation in toxicity among eight invertebrate taxa and eight sediments", *Environmental Toxicology and Chemistry*, Vol. 32/11, pp. 2495-2506, https://doi.org/10.1002/etc.2271.

Bishop, P. et al. (2012), "Animal Use and Lessons Learned in the U.S. High Production Volume Chemicals Challenge Program", *Environmental Health Perspectives*, Vol. 120/12, pp. 1631-1639, https://doi.org/10.1289/ehp.1104666.

Blackburn, K. et al. (2011), "Case studies to test: A framework for using structural, reactivity, metabolic and physicochemical similarity to evaluate the suitability of analogs for SAR-based toxicological assessments", *Regulatory Toxicology and Pharmacology*, Vol. 60/1, pp. 120-135, https://doi.org/10.1016/j.yrtph.2011.03.002.

Blackburn, K. and S. Stuard (2014), "A framework to facilitate consistent characterization of read across

uncertainty", *Regulatory Toxicology and Pharmacology*, Vol. 68/3, pp. 353-362, https://doi.org/10.1016/j.yrtph.2014.01.004.

Boots, T. et al. (2021), "Utilizing literature-based rodent toxicology data to derive potency estimates for quantitative risk assessment", *Nanotoxicology*, Vol. 15/6, pp. 740-760, https://doi.org/10.1080/17435390.2021.1918278.

Bowen, T. et al. (2023), "Simultaneously discovering the fate and biochemical effects of pharmaceuticals through untargeted metabolomics", *Nature Communications*, Vol. 14/1, https://doi.org/10.1038/s41467-023-40333-7.

Boyce, M. et al. (2023), "Identifying xenobiotic metabolites with *in silico* prediction tools and LCMS suspect screening analysis", *Frontiers in Toxicology*, Vol. 5, https://doi.org/10.3389/ftox.2023.1051483.

Boyce, M. et al. (2022), "Comparing the performance and coverage of selected *in silico* (liver) metabolism tools relative to reported studies in the literature to inform analogue selection in read-across: A case study", *Computational Toxicology*, Vol. 21, p. 100208, https://doi.org/10.1016/j.comtox.2021.100208.

Boyles, M. et al. (2015), "Multi-walled carbon nanotube induced frustrated phagocytosis, cytotoxicity and pro-inflammatory conditions in macrophages are length dependent and greater than that of asbestos", *Toxicology in Vitro*, Vol. 29/7, pp. 1513-1528, https://doi.org/10.1016/j.tiv.2015.06.012.

Braakhuis, H. et al. (2021), "An Integrated Approach to Testing and Assessment to Support Grouping and Read-Across of Nanomaterials After Inhalation Exposure", *Applied In Vitro Toxicology*, Vol. 7/3, pp. 112-128, https://doi.org/10.1089/aivt.2021.0009.

Brown, D. et al. (2001), "Size-Dependent Proinflammatory Effects of Ultrafine Polystyrene Particles: A Role for Surface Area and Oxidative Stress in the Enhanced Activity of Ultrafines", *Toxicology and Applied Pharmacology*, Vol. 175/3, pp. 191-199, https://doi.org/10.1006/taap.2001.9240.

Buick, J. et al. (2015), "Integration of metabolic activation with a predictive toxicogenomics signature to classify genotoxic versus nongenotoxic chemicals in human <scp>TK</scp>6 cells", *Environmental and Molecular Mutagenesis*, Vol. 56/6, pp. 520-534, https://doi.org/10.1002/em.21940.

Burello, E. and A. Worth (2011), "Predicting toxicity of nanoparticles", *Nature Nanotechnology*, Vol. 6/3, pp. 138-139, https://doi.org/10.1038/nnano.2011.27.

Burton, G. (2010), "Metal Bioavailability and Toxicity in Sediments", *Critical Reviews in Environmental Science and Technology*, Vol. 40/9-10, pp. 852-907, https://doi.org/10.1080/10643380802501567.

Caballero-Guzman, A. and B. Nowack (2018), "Prospective nanomaterial mass flows to the environment by life cycle stage from five applications containing CuO, DPP, FeO , CNT and SiO2", *Journal of Cleaner Production*, Vol. 203, pp. 990-1002, https://doi.org/10.1016/j.jclepro.2018.08.265.

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

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

Canada.: 124.

Canada (2010a), Phenol, (1,1-dimethylethyl)-4-methoxy-(Butylated hydroxyanisole), Screening Assessment for the Challenge. Screening assessment for the challenge : Phenol, (1,1-dimethylethyl)-4-methoxy-(Butylated hydroxyanisole) : Chemical Abstracts Service registry number 25013-16-5. : En14-360/2010E-PDF - Government of Canada Publications - Canada.ca

Canada (2010b), Decanedioic acid, bis(1,2,2,6,6-pentamethyl-4-piperidinyl)-ester. Screening Assessment for the Challenge. https://www.canada.ca/content/dam/eccc/migration/ese-ees/f638d9cf-aa79-41c6-b9b8-a27156c44c0a/batch9_41556-26-7_en.pdf

Canada (2010c), Final Second Priority Substances List Assessment Report on Three Aluminum Salts. Second Priority Substances List Assessment Report on three aluminum salts - Canada.ca

Canada (2010d), Methylium, [4-(dimethylamino)phenyl]bis[4-(ethylamino)-3-methylphenyl]-, acetate. https://www.canada.ca/en/environment-climate-change/services/evaluating-existing-substances/screening-assessment-forchallenge75.html

Casati S. et al. (2013), "EURL ECVAM Strategy for Replacement of Animal Testing for Skin Sensitisation Hazard Identification and Classification". EUR 25816 EN. https://publications.jrc.ec.europa.eu/repository/handle/JRC79446

Catlin, N. et al. (2018), “How similar is similar enough? A sufficient similarity case study with Ginkgo biloba extract”, *Food and Chemical Toxicology*, Vol. 118, pp. 328-339, https://doi.org/10.1016/j.fct.2018.05.013.

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

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

Chapkanov, A., T. Schultz and O. Mekenyan (2024), “Uncertainty in the results from oral repeated dose toxicity tests: Impact on regulatory classifications”, *Regulatory Toxicology and Pharmacology*, Vol. 146, p. 105541, https://doi.org/10.1016/j.yrtph.2023.105541.

Chapkanov, A. et al. (2023), “Estimating uncertainty in LLNA EC3 data and its impact on regulatory classifications”, *Regulatory Toxicology and Pharmacology*, Vol. 139, p. 105357, https://doi.org/10.1016/j.yrtph.2023.105357.

Chen, P. (ed.) (2020), “Metrics for graph comparison: A practitioner’s guide”, *PLOS ONE*, Vol. 15/2, p. e0228728, https://doi.org/10.1371/journal.pone.0228728.

Chilton, M. et al. (2018), “Making reliable negative predictions of human skin sensitisation using an *in silico* fragmentation approach”, *Regulatory Toxicology and Pharmacology*, Vol. 95, pp. 227-235, https://doi.org/10.1016/j.yrtph.2018.03.015.

Clemedson, C. et al. (2002), “Development of an *In Vitro*Test Battery for the Estimation of Acute Human Systemic Toxicity: An Outline of the EDIT Project”, *Alternatives to Laboratory Animals*, Vol. 30/3, pp. 313-321, https://doi.org/10.1177/026119290203000309.

Clemons, J. et al. (1997), “Derivation of 2,3,7,8-TCDD toxic equivalence factors (TEFs) for selected dioxins, furans and PCBs with rainbow trout and rat liver cell lines and the influence of exposure time”, Chemosphere, Vol. 34/5-7, pp. 1105-1119, https://doi.org/10.1016/s0045-6535(97)00412-8.

Cohen, J. et al. (2012), “Interactions of engineered nanomaterials in physiological media and

implications for *in vitro* dosimetry", *Nanotoxicology*, Vol. 7/4, pp. 417-431, https://doi.org/10.3109/17435390.2012.666576.

Comandella, D. et al. (2020), "Quality of physicochemical data on nanomaterials: an assessment of data completeness and variability", *Nanoscale*, Vol. 12/7, pp. 4695-4708, https://doi.org/10.1039/c9nr08323e.

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

Concawe (1996), Environmental risk assessment of petroleum substances: the hydrocarbon block method. Report no. 96/52, https://www.concawe.eu/wp-content/uploads/rpt_96-52-2004-01719-01-e.pdf

Concawe (2001) Classification and labelling of petroleum substances according to the EU dangerous substances directive (CONCAWE recommendations August 2001). Report no. 01/54, https://www.concawe.eu/wp-content/uploads/2002-00238-01-e.pdf

Concawe (2019), Concawe Substance Identification Group Analytical Program Report (Abridged Version). Report no. 5/19, https://www.concawe.eu/wp-content/uploads/Concawe-Substance-Identification-Group-Analytical-Program-Report-Abridged-Version.pdf

Concawe (2022), Hydrocarbon space mapping to support gas oil read-across for human health hazard assessment, https://www.concawe.eu/wp-content/uploads/Hydrocarbon-space-mapping.pdf

Creton, S. et al. (2012), "Use of toxicokinetics to support chemical evaluation: Informing high dose selection and study interpretation", *Regulatory Toxicology and Pharmacology*, Vol. 62/2, pp. 241-247, https://doi.org/10.1016/j.yrtph.2011.12.005.

Cronin, M. et al. (2022), "A scheme to evaluate structural alerts to predict toxicity – Assessing confidence by characterising uncertainties", *Regulatory Toxicology and Pharmacology*, Vol. 135, p. 105249, https://doi.org/10.1016/j.yrtph.2022.105249.

Cross, R. et al. (2022), "Assessing the similarity of nanoforms based on the biodegradation of organic surface treatment chemicals", *NanoImpact*, Vol. 26, p. 100395, https://doi.org/10.1016/j.impact.2022.100395.

Cross, R. et al. (2024), "An integrated approach to testing and assessment (IATA) to support grouping and read-across of nanomaterials in aquatic systems", *Nano Today*, Vol. 54, p. 102065, https://doi.org/10.1016/j.nantod.2023.102065.

Danzeisen, R. et al. (2022), "The underlying mode of action for lung tumors in a tiered approach to the assessment of inhaled cobalt compounds", *Regulatory Toxicology and Pharmacology*, Vol. 130, p. 105140, https://doi.org/10.1016/j.yrtph.2022.105140.

Danzeisen, R. et al. (2020), "Bioelution, Bioavailability, and Toxicity of Cobalt Compounds Correlate", *Toxicological Sciences*, Vol. 174/2, pp. 311-325, https://doi.org/10.1093/toxsci/kfz249.

Date, M. et al. (2020), "Clustering a Chemical Inventory for Safety Assessment of Fragrance Ingredients: Identifying Read-Across Analogs to Address Data Gaps", *Chemical Research in Toxicology*, Vol. 33/7, pp. 1709-1718, https://doi.org/10.1021/acs.chemrestox.9b00518.

De Abrew, K. et al. (2016), "Grouping 34 Chemicals Based on Mode of Action Using Connectivity

Mapping", *Toxicological Sciences*, Vol. 151/2, pp. 447-461, https://doi.org/10.1093/toxsci/kfw058.

De Abrew, K. et al. (2022), "A New Approach Methodology (NAM) Based Assessment of Butylated hydroxytoluene (BHT) for Endocrine Disruption Potential", *Toxicological Sciences*, Vol. 190/2, pp. 227-241, https://doi.org/10.1093/toxsci/kfac099.

De Abrew, K. et al. (2019), "Use of connectivity mapping to support read across: A deeper dive using data from 186 chemicals, 19 cell lines and 2 case studies", *Toxicology*, Vol. 423, pp. 84-94, https://doi.org/10.1016/j.tox.2019.05.008.

DeLoid, G. et al. (2014), "Estimating the effective density of engineered nanomaterials for *in vitro* dosimetry", *Nature Communications*, Vol. 5/1, https://doi.org/10.1038/ncomms4514.

Dermen, I. et al. (2022), "Estimating the reliability of simulated metabolism using documented data and theoretical knowledge. QSAR application", *Computational Toxicology*, Vol. 22, p. 100218, https://doi.org/10.1016/j.comtox.2022.100218.

Derr, R. et al. (2022), "A tiered approach to investigate the inhalation toxicity of cobalt substances. Tier 2 b: Reactive cobalt substances induce oxidative stress in ToxTracker and activate hypoxia target genes", *Regulatory Toxicology and Pharmacology*, Vol. 129, p. 105120, https://doi.org/10.1016/j.yrtph.2022.105120.

DeVito, M. et al. (2024), "The 2022 world health organization reevaluation of human and mammalian toxic equivalency factors for polychlorinated dioxins, dibenzofurans and biphenyls", *Regulatory Toxicology and Pharmacology*, Vol. 146, p. 105525, https://doi.org/10.1016/j.yrtph.2023.105525.

Di Battista, V. et al. (2024), "Similarity of multicomponent nanomaterials in a safer-by-design context: the case of core–shell quantum dots", *Environmental Science: Nano*, Vol. 11/3, pp. 924-941, https://doi.org/10.1039/d3en00338h.

Di Cristo, L. et al. (2021), "Grouping Hypotheses and an Integrated Approach to Testing and Assessment of Nanomaterials Following Oral Ingestion", *Nanomaterials*, Vol. 11/10, p. 2623, https://doi.org/10.3390/nano11102623.

Di Cristo, L. et al. (2022a), "Integrated approaches to testing and assessment for grouping nanomaterials following dermal exposure", *Nanotoxicology*, Vol. 16/3, pp. 310-332, https://doi.org/10.1080/17435390.2022.2085207.

Di Cristo et al., (2022b), "Integrated Approach to Testing and Assessment ingested silica nanomaterials showing gradual dissolution in the oro-gastrointestinal tract: grouping decisional criteria to streamline nanomaterial risk assessment". *Particle and Fibre Toxicology,* Vol. 19:/68.

Di Cristo, L. et al. (2022c), "Grouping of orally ingested silica nanomaterials via use of an integrated approach to testing and assessment to streamline risk assessment", *Particle and Fibre Toxicology*, Vol. 19/1, https://doi.org/10.1186/s12989-022-00508-4.

Diamond, G. et al. (2016), "Predicting oral relative bioavailability of arsenic in soil from *in vitro* bioaccessibility", *Journal of Toxicology and Environmental Health, Part A*, Vol. 79/4, pp. 165-173, https://doi.org/10.1080/15287394.2015.1134038.

Diderich, R. (2010), "Tools for Category Formation and Read-Across: Overview of the OECD (Q)SAR Application Toolbox", in *In Silico Toxicology*, The Royal Society of Chemistry, https://doi.org/10.1039/9781849732093-00385.

Dijkstra, J., J. Meeussen and R. Comans (2004), “Leaching of Heavy Metals from Contaminated Soils: An Experimental and Modeling Study”, *Environmental Science & Technology*, Vol. 38/16, pp. 4390-4395, https://doi.org/10.1021/es049885v.

Dimitrov, S. et al. (2015), “UVCB substances: Methodology for structural description and application to fate and hazard assessment”, *Environmental Toxicology and Chemistry*, Vol. 34/11, pp. 2450-2462, https://doi.org/10.1002/etc.3100.

Djoumbou-Feunang, Y. et al. (2019), “BioTransformer: a comprehensive computational tool for small molecule metabolism prediction and metabolite identification”, *Journal of Cheminformatics*, Vol. 11/1, https://doi.org/10.1186/s13321-018-0324-5.

Drew, N. et al. (2017), “A quantitative framework to group nanoscale and microscale particles by hazard potency to derive occupational exposure limits: Proof of concept evaluation”, *Regulatory Toxicology and Pharmacology*, Vol. 89, pp. 253-267, https://doi.org/10.1016/j.yrtph.2017.08.003.

Driscoll, K. (2022), “Review of Lung Particle Overload, Rat Lung Cancer, and the Conclusions of the Edinburgh Expert Panel—It’s Time to Revisit Cancer Hazard Classifications for Titanium Dioxide and Carbon Black”, *Frontiers in Public Health*, Vol. 10, https://doi.org/10.3389/fpubh.2022.907318.

Duffin, R. et al. (2002), “The Importance of Surface Area and Specific Reactivity in the Acute Pulmonary Inflammatory Response to Particles”, The Annals of Occupational Hygiene, https://doi.org/10.1093/annhyg/46.suppl_1.242.

Duffin, R. et al. (2007), “Proinflammogenic Effects of Low-Toxicity and Metal Nanoparticles *In Vivo* and *In Vitro*: Highlighting the Role of Particle Surface Area and Surface Reactivity”, *Inhalation Toxicology*, Vol. 19/10, pp. 849-856, https://doi.org/10.1080/08958370701479323.

Dutton, M. et al. (2019), “Towards an exposure narrative for metals and arsenic in historically contaminated Ni refinery soils: Relationships between speciation, bioavailability, and bioaccessibility”, *Science of The Total Environment*, Vol. 686, pp. 805-818, https://doi.org/10.1016/j.scitotenv.2019.05.164.

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

EC (2008), Regulation (EC) No 1272/2008 of the European Parliament and of the Council of 16 December 2008 on classification, labelling and packaging of substances and mixtures, amending and repealing the Dangerous Substances Directive 67/548/EEC and the Dangerous Preparations Directive 1999/45/EC, and amending Regulation (EC) No 1907/2006; amended by Regulation (EU) 2024/2865 of the European Parliament and of the Council of 23 October 2024

EC (2017), EC Joint Research Centre, "Evaluation of the availability and applicability of computational approaches in the safety assessment of nanomaterials: final report of the Nanocomput project", Publications Office, https://data.europa.eu/doi/10.2760/248139

EC (2022), Commission Recommendation of 10 June 2022 on the definition of nanomaterial (2022/C 229/01), Off. J. Eur. Union C 229/1-5, 14.06.2022.

ECB (2005), Scoping study on the development of a technical guidance document on information

requirements on intrinsic properties of substances (RIP 3.3-1), CEFIC, DK-EPA, Environmental Agency of Wales and England, ECETOC, INERIS, KemI and TNO. European Chemicals Bureau, Joint Research Centre, European Commission, Ispra, Italy.

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

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

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

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

ECETOC Nanoapp. Available at https://www.ecetoc.org/tools/nanoapp/

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

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

ECCC (2016), Screening assessment, Aromatic Azo and Benzidine-based Substance Grouping. https://www.canada.ca/en/environment-climate-change/services/evaluating-existing-substances/screening-assessment11.html

ECCC (2017), Screening assessment, Selenium and its compounds. https://www.canada.ca/en/environment-climate-change/services/evaluating-existing-substances/screening-assessment-selenium.html

EEC (1993) No. 793/93 of 23 March 1993 on the evaluation and control of the risks of existing substances. http://ec.europa.eu/environment/chemicals/exist_subst/pdf/793_93_ec.pdf

ECHA (2008a), "Guidance on information requirements and chemical safety assessment, Chapter R.6: QSARs and grouping of chemicals" https://echa.europa.eu/documents/10162/17224/information_requirements_r6_en.pdf/77f49f81-b76d-40ab-8513-4f3a533b6ac9

ECHA (2008b), "Guidance on information requirements and chemical safety assessment Appendix R.7.13-2: Environmental risk assessment for metals and metal compounds", https://echa.europa.eu/documents/10162/17224/information_requirements_r7_13_2_en.pdf/0497e68d-4bb5-4b12-a4db-52ce0c1bc237.

ECHA (2016), "How to use alternatives to animal testing to fulfil your information requirements for REACH registration", ECHA-16-B-25-EN, ECHA, Helsinki, https://doi.org/10.2823/194297.

ECHA (2017a), "Guidance on information requirements and chemical safety assessment Appendix R.6-1 for nanomaterials applicable to the Guidance on QSARs and Grouping of Chemicals", ECHA-17-G-17-EN, ECHA, Helsinki.

ECHA (2017b), "Read-Across Assessment Framework (RAAF)", ECHA-17-R-01-EN, ECHA, Helsinki, https://doi.org/10.2823/619212.

ECHA (2017c) “Read-Across Assessment Framework (RAAF) - considerations on multi-constituent substances and UVCBs”, ECHA-17-R-04-EN, ECHA, Helsinki, https://doi.org/10.2823/794394.

ECHA (2017d), "The use of alternatives to testing on animals for the REACH Regulation - Third report under Article 117(3) of the REACH Regulation", ECHA-17-R-02-EN, ECHA, Helsinki, https://doi.org/10.2823/023078.

ECHA (2019), “Guidance on information requirements and chemical safety assessment Appendix R.6-1 for nanoforms applicable to the Guidance on QSARs and Grouping of Chemicals", version 2.0, ECHA-19-H-15-EN, ECHA, Helsinki, https://doi.org/10.2823/273911.

ECHA (2020), "The use of alternatives to testing on animals for the REACH Regulation - Fourth report (2020) under Article 117(3) of the REACH Regulation”, ECHA-20-R-08-EN, ECHA, Helsinki, https://doi.org/10.2823/092305.

ECHA (2021), "Assessment of Regulatory Needs (ARN). Simple Vanadium compounds". https://www.echa.europa.eu/documents/10162/62106e4c-e830-543e-5176-16591be1ea8a

ECHA (2022a), "Assessment of Regulatory Needs (ARN). Simple Lithium compounds". https://www.echa.europa.eu/documents/10162/b43e6224-a15c-c83d-6969-935e41748dad

ECHA (2022b), "Advice on using read-across for UVCB substances" https://echa.europa.eu/documents/10162/11395738/advice_uvcb_read-across_en.pdf/ac1f64a6-9ee5-441e-cf1c-92914b843b4e?t=1651665130365

ECHA (2022c), “Appendix for nanoforms applicable to the Guidance on Registration and Substance Identification”, ECHA-21-G-06-EN, ECHA, Helsinki, https://doi.org/10.2823/40.

ECHA (2023a), "The use of alternatives to testing on animals for the REACH Regulation- Fifth report under Article 117(3) of the REACH Regulation", ECHA-23-R-07-EN, ECHA, Helsinki, https://doi.org/10.2823/805454.

ECHA (2023b), "Guidance for identification and naming of substances under REACH and CLP", ECHA-23-H-07-EN, ECHA, Helsinki, https://doi.org/10.2823/87416.

ECHA (2024a), IUCLID 6 User Manual. https://iuclid6.echa.europa.eu/documents/1387205/1809908/iuclid_user_manual_en.pdf

ECHA (2024b), “Guidance on the Application of the CLP Criteria - Part 4: Environmental hazards and Part 5: Additional Hazards. Guidance to Regulation (EC) No 1272/2008 on classification, labelling and packaging (CLP) of substances and mixtures”, ECHA-24-G-05-EN, ECHA, Helsinki, https://doi.org/10.2823/4895485.

EFSA (2018a), "Guidance on Uncertainty Analysis in Scientific Assessments", EFSA Journal 2018;16(1):5123, 39 pp. https://doi.org/10.2903/j.efsa.2018.5123

EFSA (2018b), "Principles and methods behind EFSA’s Guidance on Uncertainty Analysis in Scientific Assessment". EFSA Journal 2018;16(1):5122, 282 pp. https://doi.org/10.2903/j.efsa.2018.5122

EFSA (2019), "Dietary reference values for sodium". Adopted: 3 July 2019. doi:10.2903/j.efsa.2019.5778

EFSA Scientific Committee (2021), “Guidance on risk assessment of nanomaterials to be applied in the food and feed chain: human and animal health”, *EFSA Journal*, Vol. 19/8, https://doi.org/10.2903/j.efsa.2021.6768.

Elphick, J., K. Bergh and H. Bailey (2011), “Chronic toxicity of chloride to freshwater species: Effects of hardness and implications for water quality guidelines”, *Environmental Toxicology and Chemistry*,

Vol. 30/1, pp. 239-246, https://doi.org/10.1002/etc.365.

Enoch, S. (2009), “Chemical Category Formation and Read-Across for the Prediction of Toxicity”, in *Challenges and Advances in Computational Chemistry and Physics, Recent Advances in QSAR Studies*, Springer Netherlands, Dordrecht, https://doi.org/10.1007/978-1-4020-9783-6_7.

Enoch, S. and M. Cronin (2010), “A review of the electrophilic reaction chemistry involved in covalent DNA binding”, *Critical Reviews in Toxicology*, Vol. 40/8, pp. 728-748, https://doi.org/10.3109/10408444.2010.494175.

Enoch, S. et al. (2011), “A review of the electrophilic reaction chemistry involved in covalent protein binding relevant to toxicity”, *Critical Reviews in Toxicology*, Vol. 41/9, pp. 783-802, https://doi.org/10.3109/10408444.2011.598141.

Enoch, S. et al. (2023), “Metabolism-based category formation for the prioritisation of genotoxicity hazard assessment for plant protection product residues (part 3): Strobilurins”, *Regulatory Toxicology and Pharmacology*, Vol. 144, p. 105484, https://doi.org/10.1016/j.yrtph.2023.105484.

Enoch, S. et al. (2022), “Sub-structure-based category formation for the prioritisation of genotoxicity hazard assessment for pesticide residues: Sulphonyl ureas”, *Regulatory Toxicology and Pharmacology*, Vol. 129, p. 105115, https://doi.org/10.1016/j.yrtph.2022.105115.

Enoch, S. et al. (2012), “Development of Mechanism-Based Structural Alerts for Respiratory Sensitization Hazard Identification”, *Chemical Research in Toxicology*, Vol. 25/11, pp. 2490-2498, https://doi.org/10.1021/tx3003092.

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

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

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

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

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

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

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

EPA (2004b), "High Production Volume (HPV) Challenge Robust Summaries & Test Plans: Terpenoid Primary Alcohols and Related Esters Category". http://www.epa.gov/hpv/pubs/summaries/terprial/c12965tc.htm

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

EPA (2005), "High Production Volume (HPV) Challenge. Robust Summaries & Test Plans: Cinnamyl

Derivatives". http://www.epa.gov/hpv/pubs/summaries/cinna/c12912tc.htm

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

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

EPA (2010a), "High production volume (HPV) challenge: Robust summaries & test plans: Aluminum alkoxides category". https://hero.epa.gov/hero/index.cfm/reference/details/reference_id/1508266

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

EPA (2014), "Screening-Level Hazard Characterization Heavy Fuel Oils Category". https://nepis.epa.gov/Exe/ZyPURL.cgi?Dockey=P10104W8.txt

EPA (2021), "Provisional Peer-Reviewed Toxicity Values for 3,4-Toluenediamine". https://hhpprtv.ornl.gov/issue_papers/Toluenediamine34.pdf (CASRN 496-72-0)

Erickson, R. et al. (2016), “The acute toxicity of major ion salts to *Ceriodaphnia dubia*. II. Empirical relationships in binary salt mixtures”, *Environmental Toxicology and Chemistry*, Vol. 36/6, pp. 1525-1537, https://doi.org/10.1002/etc.3669.

Escher, S. et al. (2019), “Towards grouping concepts based on new approach methodologies in chemical hazard assessment: the read-across approach of the EU-ToxRisk project”, *Archives of Toxicology*, Vol. 93/12, pp. 3643-3667, https://doi.org/10.1007/s00204-019-02591-7.

Escher, S. et al. (2022), “Integrate mechanistic evidence from new approach methodologies (NAMs) into a read-across assessment to characterise trends in shared mode of action”, *Toxicology in Vitro*, Vol. 79, p. 105269, https://doi.org/10.1016/j.tiv.2021.105269.

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

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

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

EU-ToxRisk (2018), "Recommendations of the EU-ToxRisk Regulatory Advisory Board (RAB) on how to document case studies for regulatory evaluation".

Fabjan, E. et al. (2006), “A Category Approach for Reproductive Effects of Phthalates”, *Critical Reviews in Toxicology*, Vol. 36/9, pp. 695-726, https://doi.org/10.1080/10408440600894914.

Fernández-Cruz, M. et al. (2018), “Quality evaluation of human and environmental toxicity studies performed with nanomaterials – the GUIDEnano approach”, *Environmental Science: Nano*, Vol. 5/2, pp. 381-397, https://doi.org/10.1039/c7en00716g.

Firman, J. et al. (2022), “Construction of an *In Silico* Structural Profiling Tool Facilitating Mechanistically

Grounded Classification of Aquatic Toxicants", Environmental Science & Technology, Vol. 56/24, pp. 17805-17814, https://doi.org/10.1021/acs.est.2c03736.

Gadaleta, D. (2020), "Automated Integration of Structural, Biological and Metabolic Similarities to Sustain Read-Across", *ALTEX*, https://doi.org/10.14573/altex.2002281.

Gerloff, K. et al. (2017), "The Adverse Outcome Pathway approach in nanotoxicology", *Computational Toxicology*, Vol. 1, pp. 3-11, https://doi.org/10.1016/j.comtox.2016.07.001.

Giusti, A. et al. (2019), "Nanomaterial grouping: Existing approaches and future recommendations", *NanoImpact*, Vol. 16, p. 100182, https://doi.org/10.1016/j.impact.2019.100182.

Goyer, R.A. (1996), Chapter 23 toxic effects of metals. In: Klaassen, C.D. (Ed.), Cassarett and Doull's Toxicology: the Basic Science of Poisons, fifth ed. McGraw-Hill, New York, pp. 691–736.

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

Grimm, F. et al. (2016), "A chemical–biological similarity-based grouping of complex substances as a prototype approach for evaluating chemical alternatives", *Green Chemistry*, Vol. 18/16, pp. 4407-4419, https://doi.org/10.1039/c6gc01147k.

Groff, L. et al. (2024), "MetSim: Integrated Programmatic Access and Pathway Management for Xenobiotic Metabolism Simulators", *Chemical Research in Toxicology*, Vol. 37/5, pp. 685-697, https://doi.org/10.1021/acs.chemrestox.3c00398.

Grulke, C. et al. (2019), "EPA's DSSTox database: History of development of a curated chemistry resource supporting computational toxicology research", *Computational Toxicology*, Vol. 12, p. 100096, https://doi.org/10.1016/j.comtox.2019.100096.

Gruszczynska, H. et al. (2024), "Multi-omics bioactivity profile-based chemical grouping and read-across: a case study with *Daphnia magna* and azo dyes", *Archives of Toxicology*, Vol. 98/8, pp. 2577-2588, https://doi.org/10.1007/s00204-024-03759-6.

Hanway, R.H. and P.F. Evans (2000), "Read-across of toxicological data in the notification of new chemicals", *Toxicology Letters*, Vol. 116/1, pp. 61.

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

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

Hartmann, N. et al. (2017), "NanoCRED: A transparent framework to assess the regulatory adequacy of ecotoxicity data for nanomaterials – Relevance and reliability revisited", *NanoImpact*, Vol. 6, pp. 81-89, https://doi.org/10.1016/j.impact.2017.03.004.

Heim, K. et al. (2020), "Bioaccessibility of nickel and cobalt in synthetic gastric and lung fluids and its potential use in alloy classification", *Regulatory Toxicology and Pharmacology*, Vol.110, pp. 104549. https://doi.org/10.1016/j.yrtph.2019.104549

Helman, G., I. Shah and G. Patlewicz (2018), "Extending the Generalised Read-Across approach (GenRA): A systematic analysis of the impact of physicochemical property information on read-across performance", *Computational Toxicology*, Vol. 8, pp. 34-50,

https://doi.org/10.1016/j.comtox.2018.07.001.

Henderson, R. et al. (2012), “Oral bioaccessibility testing and read-across hazard assessment of nickel compounds”, *Regulatory Toxicology and Pharmacology*, Vol. 63/1, pp. 20-28, https://doi.org/10.1016/j.yrtph.2012.02.005.

Hilton, G. et al. (2022), “Rethinking chronic toxicity and carcinogenicity assessment for agrochemicals project (ReCAAP): A reporting framework to support a weight of evidence safety assessment without long-term rodent bioassays”, *Regulatory Toxicology and Pharmacology*, Vol. 131, p. 105160, https://doi.org/10.1016/j.yrtph.2022.105160.

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

House, J. (2022), “Grouping of UVCB substances with dose-response transcriptomics data from human cell-based assays”, *ALTEX*, https://doi.org/10.14573/altex.2107051.

House, J. (2020), “Grouping of UVCB substances with new approach methodologies (NAMs) data”, *ALTEX*, https://doi.org/10.14573/altex.2006262.

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 (HSPA), Brussels,

HSPA (2019), "Substance identification and naming convention for hydrocarbon solvents" Hydrocarbon Solvents Producers Association (HSPA), Brussels, https://www.esig.org/wp-content/uploads/2019/12/201910_HSPA-naming-convention-_October-2019.pdf

Hund-Rinke et al. (2017). "Considerations about the relationship of nanomaterial’s physical-chemical properties and aquatic toxicity for the purpose of grouping". https://www.umweltbundesamt.de/publikationen/considerations-about-the-relationship-of

IARC (1997), "Monographs on the Evaluation of Carcinogenic Risks to Humans, Silica, Some Silicates, Coal Dist and para-Aramid Fibrils"., Report 68.

IARC (2002), " IARC monographs on the evaluation of carcinogenic risks to humans", Man-made vitreous fibres, Report Volume 81. ISBN-13 978-92-832-1281-2.

ICRP (1994). "Human Respiratory Tract Model for Radiological Protection"c. ICRP Publication 66. Annals of the ICRP, 24, 698-703.

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

Ingre-Khans, E. et al. (2019), “Reliability and relevance evaluations of REACH data”, *Toxicology Research*, Vol. 8/1, pp. 46-56, https://doi.org/10.1039/c8tx00216a.

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

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

globally-harmonized-system-ghs-criteria-petroleum-substances.

Irwan, J. et al. (2024), “Read-Across Application for Food or Feed Ingredients”, *EFSA Supporting Publications*, Vol. 21/7, https://doi.org/10.2903/sp.efsa.2024.en-8811.

Jäckh, R. (2007), Investigations of read-across & grouping, including mechanistic considerations and rationales for outliers. In: A Compendium of Case Studies that helped to shape the REACH Guidance on Chemical Categories and Read Across, Worth A. & Patlewicz G. (Eds.), EUR report no 22481 EN. https://op.europa.eu/en/publication-detail/-/publication/3339157f-7e08-4ba1-9b9b-238b42a7f402

Janer, G. et al. (2021), “Creating sets of similar nanoforms with the ECETOC NanoApp: real-life case studies”, *Nanotoxicology*, Vol. 15/8, pp. 1016-1034, https://doi.org/10.1080/17435390.2021.1946186.

Janer, G., R. Landsiedel and W. Wohlleben (2020), “Rationale and decision rules behind the ECETOC NanoApp to support registration of sets of similar nanoforms within REACH”, *Nanotoxicology*, Vol. 15/2, pp. 145-166, https://doi.org/10.1080/17435390.2020.1842933.

Jeliazkova, N. et al. (2021), “Towards FAIR nanosafety data”, *Nature Nanotechnology*, Vol. 16/6, pp. 644-654, https://doi.org/10.1038/s41565-021-00911-6.

Jeliazkova, N. et al. (2022a), “How can we justify grouping of nanoforms for hazard assessment? Concepts and tools to quantify similarity”, *NanoImpact*, Vol. 25, p. 100366, https://doi.org/10.1016/j.impact.2021.100366.

Jeliazkova, N. et al. (2022b), “Possibilities to group nanomaterials across different substances – A case study on organic pigments”, *NanoImpact*, Vol. 26, p. 100391, https://doi.org/10.1016/j.impact.2022.100391.

Judson, R. et al. (2015), “Integrated Model of Chemical Perturbations of a Biological Pathway Using 18 *In Vitro* High-Throughput Screening Assays for the Estrogen Receptor”, *Toxicological Sciences*, Vol. 148/1, pp. 137-154, https://doi.org/10.1093/toxsci/kfv168.

Keller, J. et al. (2021), “Variation in dissolution behavior among different nanoforms and its implication for grouping approaches in inhalation toxicity”, *NanoImpact*, Vol. 23, p. 100341, https://doi.org/10.1016/j.impact.2021.100341.

Kennedy, A. et al. (2017), “Assessing nanomaterial exposures in aquatic ecotoxicological testing: Framework and case studies based on dispersion and dissolution”, *Nanotoxicology*, Vol. 11/4, pp. 546-557, https://doi.org/10.1080/17435390.2017.1317863.

Kessler, A. et al. (2023), “Unravelling the Mechanistic Understanding of Metal Nanoparticle-Induced Reactive Oxygen Species Formation: Insights from a Cu Nanoparticle Study”, *Chemical Research in Toxicology*, Vol. 36/12, pp. 1891-1900, https://doi.org/10.1021/acs.chemrestox.3c00177.

Kirkland, D. et al. (2015), “New investigations into the genotoxicity of cobalt compounds and their impact on overall assessment of genotoxic risk”, *Regulatory Toxicology and Pharmacology*, Vol. 73/1, pp. 311-338, https://doi.org/10.1016/j.yrtph.2015.07.016.

Klimisch, H., M. Andreae and U. Tillmann (1997), “A Systematic Approach for Evaluating the Quality of Experimental Toxicological and Ecotoxicological Data”, *Regulatory Toxicology and Pharmacology*, Vol. 25/1, pp. 1-5, https://doi.org/10.1006/rtph.1996.1076.

Koval, L. et al. (2022), “Wildfire Variable Toxicity: Identifying Biomass Smoke Exposure Groupings through Transcriptomic Similarity Scoring”, *Environmental Science & Technology*, Vol. 56/23, pp. 17131-17142, https://doi.org/10.1021/acs.est.2c06043.

Kreyling, W. et al. (2013), “Air–Blood Barrier Translocation of Tracheally Instilled Gold Nanoparticles Inversely Depends on Particle Size”, *ACS Nano*, Vol. 8/1, pp. 222-233, https://doi.org/10.1021/nn403256v.

Kreyling, W. et al. (2017), “Quantitative biokinetics of titanium dioxide nanoparticles after intravenous injection in rats: Part 1”, *Nanotoxicology*, Vol. 11/4, pp. 434-442, https://doi.org/10.1080/17435390.2017.1306892.

Kuempel, E. et al. (2012), “Development of risk-based nanomaterial groups for occupational exposure control”, *Journal of Nanoparticle Research*, Vol. 14/9, https://doi.org/10.1007/s11051-012-1029-8.

Kuseva, C. et al. (2021), “Criteria for quantitative assessment of metabolic similarity between chemicals. II. Application to human health endpoints”, *Computational Toxicology*, Vol. 19, p. 100173, https://doi.org/10.1016/j.comtox.2021.100173.

Kutarna, S. et al. (2023), “Widespread presence of chlorinated paraffins in consumer products”, *Environmental Science: Processes & Impacts*, Vol. 25/5, pp. 893-900, https://doi.org/10.1039/d2em00494a.

Lai, A. et al. (2022), “The Next Frontier of Environmental Unknowns: Substances of Unknown or Variable Composition, Complex Reaction Products, or Biological Materials (UVCBs)”, *Environmental Science & Technology*, Vol. 56/12, pp. 7448-7466, https://doi.org/10.1021/acs.est.2c00321.

Lamon, L. et al. (2019), “Physiologically based mathematical models of nanomaterials for regulatory toxicology: A review”, *Computational Toxicology*, Vol. 9, pp. 133-142, https://doi.org/10.1016/j.comtox.2018.10.002.

Lamon, L. et al. (2019), “Computational models for the assessment of manufactured nanomaterials: Development of model reporting standards and mapping of the model landscape”, *Computational Toxicology*, Vol. 9, pp. 143-151, https://doi.org/10.1016/j.comtox.2018.12.002.

Lester, C. et al. (2018), “Structure activity relationship (SAR) toxicological assessments: The role of expert judgment”, *Regulatory Toxicology and Pharmacology*, Vol. 92, pp. 390-406, https://doi.org/10.1016/j.yrtph.2017.12.026.

Lester, C. and G. Yan (2021), “A matched molecular pair (MMP) approach for selecting analogs suitable for structure activity relationship (SAR)-based read across”, *Regulatory Toxicology and Pharmacology*, Vol. 124, p. 104966, https://doi.org/10.1016/j.yrtph.2021.104966.

Lester, C. et al. (2023), “Quantifying Analogue Suitability for SAR-Based Read-Across Toxicological Assessment”, *Chemical Research in Toxicology*, Vol. 36/2, pp. 230-242, https://doi.org/10.1021/acs.chemrestox.2c00311.

Li, J. et al. (2022), “Nano-QSAR modeling for predicting the cytotoxicity of metallic and metal oxide nanoparticles: A review”, *Ecotoxicology and Environmental Safety*, Vol. 243, p. 113955, https://doi.org/10.1016/j.ecoenv.2022.113955.

Li, S. and L. Huang (2009), “Nanoparticles evading the reticuloendothelial system: Role of the supported bilayer”, *Biochimica et Biophysica Acta (BBA) - Biomembranes*, Vol. 1788/10, pp. 2259-2266, https://doi.org/10.1016/j.bbamem.2009.06.022.

Little, S. et al. (2021), “Acute waterborne and chronic sediment toxicity of silver and titanium dioxide nanomaterials towards the oligochaete, Lumbriculus variegatus”, *NanoImpact*, Vol. 21, p. 100291, https://doi.org/10.1016/j.impact.2020.100291.

Lizarraga, L. et al. (2019), “A case study on the application of an expert-driven read-across approach in support of quantitative risk assessment of p,p’-dichlorodiphenyldichloroethane”, *Regulatory Toxicology and Pharmacology*, Vol. 103, pp. 301-313, https://doi.org/10.1016/j.yrtph.2019.02.010.

Loosli, F. et al. (2022), “Refinement of the selection of physicochemical properties for grouping and read-across of nanoforms”, *NanoImpact*, Vol. 25, p. 100375, https://doi.org/10.1016/j.impact.2021.100375.

Lowry, G. et al. (2012), “Transformations of Nanomaterials in the Environment”, *Environmental Science & Technology*, Vol. 46/13, pp. 6893-6899, https://doi.org/10.1021/es300839e.

Lubinski, L. et al. (2013), “Evaluation criteria for the quality of published experimental data on nanomaterials and their usefulness for QSAR modelling”, *SAR and QSAR in Environmental Research*, Vol. 24/12, pp. 995-1008, https://doi.org/10.1080/1062936x.2013.840679.

Madden, J. et al. (2020), “A Review of *In Silico*Tools as Alternatives to Animal Testing: Principles, Resources and Applications”, *Alternatives to Laboratory Animals*, Vol. 48/4, pp. 146-172, https://doi.org/10.1177/0261192920965977.

Ma, G. et al. (2021), “Deep graph similarity learning: a survey”, *Data Mining and Knowledge Discovery*, Vol. 35/3, pp. 688-725, https://doi.org/10.1007/s10618-020-00733-5.

Mansouri, K. et al. (2016), “CERAPP: Collaborative Estrogen Receptor Activity Prediction Project”, *Environmental Health Perspectives*, Vol. 124/7, pp. 1023-1033, https://doi.org/10.1289/ehp.1510267.

Mansouri, K. et al. (2018), “OPERA models for predicting physicochemical properties and environmental fate endpoints”, *Journal of Cheminformatics*, Vol. 10/1, https://doi.org/10.1186/s13321-018-0263-1.

Mansouri, K. et al. (2020), “CoMPARA: Collaborative Modeling Project for Androgen Receptor Activity”, *Environmental Health Perspectives*, Vol. 128/2, https://doi.org/10.1289/ehp5580.

Marchese Robinson, R. et al. (2016), “How should the completeness and quality of curated nanomaterial data be evaluated?”, *Nanoscale*, Vol. 8/19, pp. 9919-9943, https://doi.org/10.1039/c5nr08944a.

Marshall, S. et al. (2013), “An Empirical Approach to Sufficient Similarity: Combining Exposure Data and Mixtures Toxicology Data”, *Risk Analysis*, Vol. 33/9, pp. 1582-1595, https://doi.org/10.1111/risa.12015.

McCallum, M. et al. (2013), “High-Throughput Identification of Promiscuous Inhibitors from Screening Libraries with the Use of a Thiol-Containing Fluorescent Probe”, *SLAS Discovery*, Vol. 18/6, pp. 705-713, https://doi.org/10.1177/1087057113476090.

Mckee, R., M. Adenuga and J. Carrillo (2015), “Characterization of the toxicological hazards of hydrocarbon solvents”, *Critical Reviews in Toxicology*, Vol. 45/4, pp. 273-365, https://doi.org/10.3109/10408444.2015.1016216.

Mech, A. et al. (2018), “Insights into possibilities for grouping and read-across for nanomaterials in EU

chemicals legislation", *Nanotoxicology*, Vol. 13/1, pp. 119-141, https://doi.org/10.1080/17435390.2018.1513092.

Mekenyan, O. et al. (2004), "Identification of the Structural Requirements for Mutagenicity by Incorporating Molecular Flexibility and Metabolic Activation of Chemicals I: TA100 Model", *Chemical Research in Toxicology*, Vol. 17/6, pp. 753-766, https://doi.org/10.1021/tx030049t.

Mekenyan, O. et al. (2010), "Use of Genotoxicity Information in the Development of Integrated Testing Strategies (ITS) for Skin Sensitization", *Chemical Research in Toxicology*, Vol. 23/10, pp. 1519-1540, https://doi.org/10.1021/tx100161j.

Mekenyan, O. et al. (2012), "Investigating the Relationship between *in Vitro–in Vivo* Genotoxicity: Derivation of Mechanistic QSAR Models for *in Vivo* Liver Genotoxicity and *in Vivo* Bone Marrow Micronucleus Formation Which Encompass Metabolism", *Chemical Research in Toxicology*, Vol. 25/2, pp. 277-296, https://doi.org/10.1021/tx200547s.

Mekenyan, O. et al. (2007), "Identifying the Structural Requirements for Chromosomal Aberration by Incorporating Molecular Flexibility and Metabolic Activation of Chemicals", *Chemical Research in Toxicology*, Vol. 20/12, pp. 1927-1941, https://doi.org/10.1021/tx700249q.

Moermond, C. et al. (2015), "CRED: Criteria for reporting and evaluating ecotoxicity data", *Environmental Toxicology and Chemistry*, Vol. 35/5, pp. 1297-1309, https://doi.org/10.1002/etc.3259.

Molander, L. et al. (2014), "Science in Risk Assessment and Policy (SciRAP): An Online Resource for Evaluating and Reporting *In Vivo* (Eco)Toxicity Studies", *Human and Ecological Risk Assessment: An International Journal*, Vol. 21/3, pp. 753-762, https://doi.org/10.1080/10807039.2014.928104.

Mount, D. et al. (1997), "Statistical models to predict the toxicity of major ions to *Ceriodaphnia dubia, Daphnia magna and Pimephales promelas* (fathead minnows)", *Environmental Toxicology and Chemistry*, Vol. 16/10, pp. 2009-2019, https://doi.org/10.1002/etc.5620161005.

Moustakas, H. et al. (2022), "An End Point-Specific Framework for Read-Across Analog Selection for Human Health Effects", *Chemical Research in Toxicology*, Vol. 35/12, pp. 2324-2334, https://doi.org/10.1021/acs.chemrestox.2c00286.

Muir, D. et al. (1992), "Dietary 2,3,7,8-tetrachlorodibenzofuran in rainbow trout: Accumulation, disposition, and hepatic mixed-function oxidase enzyme induction", *Toxicology and Applied Pharmacology*, Vol. 117/1, pp. 65-74, https://doi.org/10.1016/0041-008x(92)90218-h.

Murphy, F. et al. (2023), "How to formulate hypotheses and IATA to support grouping and read-across of nanoforms", *ALTEX*, Vol. 40, pp. 125-140. https://doi.org/10.14573/altex.2203241.

Murphy, F. et al. (2021), "An integrated approach to testing and assessment of high aspect ratio nanomaterials and its application for grouping based on a common mesothelioma hazard", *NanoImpact*, Vol. 22, p. 100314, https://doi.org/10.1016/j.impact.2021.100314.

Murphy, F. et al. (2022), "Grouping MWCNTs based on their similar potential to cause pulmonary hazard after inhalation: a case-study", *Particle and Fibre Toxicology*, Vol. 19/1, https://doi.org/10.1186/s12989-022-00487-6.

Murray, F. et al. (2013), "The relationship between developmental toxicity and aromatic-ring class profile of high-boiling petroleum substances", *Regulatory Toxicology and Pharmacology*, Vol. 67/2, pp. S46-S59, https://doi.org/10.1016/j.yrtph.2013.05.003.

Myatt, G. et al. (2018), “*In silico* toxicology protocols”, *Regulatory Toxicology and Pharmacology*, Vol. 96, pp. 1-17, https://doi.org/10.1016/j.yrtph.2018.04.014.

Myatt, G. et al. (2022), “Implementation of *in silico* toxicology protocols within a visual and interactive hazard assessment platform”, *Computational Toxicology*, Vol. 21, p. 100201, https://doi.org/10.1016/j.comtox.2021.100201.

NAFTA (2012), "Technical working group on pesticides (TWG) (Quantitative) structure activity relationship [(Q)SAR] guidance document", http://www.epa.gov/oppfead1/international/naftatwg/guidance/qsar-guidance.pdf

National Library of Medicine website; https://pubchem.ncbi.nlm.nih.gov/

Nederstigt, T. et al. (2022a), “Correlation analysis of single- and multigenerational endpoints in Daphnia magna toxicity tests: A case-study using TiO2 nanoparticles”, *Ecotoxicology and Environmental Safety*, Vol. 241, p. 113792, https://doi.org/10.1016/j.ecoenv.2022.113792.

Nederstigt, T. et al. (2022b), “Impacts of a novel controlled-release TiO2-coated (nano-) formulation of carbendazim and its constituents on freshwater macroinvertebrate communities”, *Science of The Total Environment*, Vol. 838, p. 156554, https://doi.org/10.1016/j.scitotenv.2022.156554.

Nel, A. et al. (2013), “A Multi-Stakeholder Perspective on the Use of Alternative Test Strategies for Nanomaterial Safety Assessment”, *ACS Nano*, Vol. 7/8, pp. 6422-6433, https://doi.org/10.1021/nn4037927.

Nelms, M. et al. (2019), “Comparing and contrasting the coverage of publicly available structural alerts for protein binding”, *Computational Toxicology*, Vol. 12, p. 100100, https://doi.org/10.1016/j.comtox.2019.100100.

Netzeva, T., M. Pavan and A. Worth (2008), “Review of (Quantitative) Structure–Activity Relationships for Acute Aquatic Toxicity”, *QSAR & Combinatorial Science*, Vol. 27/1, pp. 77-90, https://doi.org/10.1002/qsar.200710099.

NIOSH (2019), "Technical report: the NIOSH occupational exposure banding process for chemical risk management". U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health, DHHS (NIOSH) Publication No. 2019-132. https://doi.org/10.26616/NIOSHPUB2019132.

NIOSH (2021), "NIOSH Approaches to Developing Occupational Exposure Limits or Bands for Engineered Nanomaterials - External Review Draft". https://www.federalregister.gov/documents/2021/07/13/2021-14801/draft-approaches-to-developing-occupational-exposure-limits-or-bands-for-engineered-nanomaterials.

Niyogi, S. and C. Wood (2004), “Biotic Ligand Model, a Flexible Tool for Developing Site-Specific Water Quality Guidelines for Metals”, *Environmental Science & Technology*, Vol. 38/23, pp. 6177-6192, https://doi.org/10.1021/es0496524.

Oberdörster, G. et al. (2005), “Principles for characterizing the potential human health effects from exposure to nanomaterials: elements of a screening strategy”, *Particle and Fibre Toxicology*, Vol. 2/1, https://doi.org/10.1186/1743-8977-2-8.

OECD (2000a), Section 3.3 of the OECD Manual for the Assessment of Chemicals, Guidance for the use of Structured Activity Relationships (SAR) in the HPV Chemicals Programme, OECD, Paris. https://web-archive.oecd.org/2012-06-14/95771-47983412.pdf

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

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

OECD (2001b), Guidance Document on Transformation/Dissolution of Metals and Metal Compounds in Aqueous Media, OECD Series on Testing and Assessment, No. 29, OECD Publishing, Paris, https://doi.org/10.1787/9789264078451-en.

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

OECD (2004a), OECD Principles for the Validation, for Regulatory Purposes, of (Quantitative) Structure- Activity Relationship Models. https://www.oecd.org/content/dam/oecd/en/topics/policy-sub-issues/assessment-of-chemicals/oecd-principles-for-the-validation-for-regulatory-purposes-of-quantitative-structure-activity-relationship-models.pdf

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

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

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

OECD (2004e), Test No. 202: Daphnia sp. Acute Immobilisation Test, OECD Guidelines for the Testing of Chemicals, Section 2, OECD Publishing, Paris, https://doi.org/10.1787/9789264069947-en.

OECD (2004f), Report from the Expert Group on (Quantitative) Structure-Activity Relationships [(Q)SARs] on the Principles for the Validation of (Q)SARs. ENV/JM/MONO(2004)24. OECD, Paris.

OECD (2004g), OECD principles for the Validation, for Regulatory Purposes, of (Q)SAR Models, ENV/JM/MONO(2004)24, No. 49, OECD Publishing, Paris.

OECD (2005a), Section 3.1 of the OECD Manual for the Assessment of Chemicals, Guidance for Determining the Quality of Data for the SIDS Dossier (Reliability, Relevance and Adequacy). OECD, Paris. https://web-archive.oecd.org/2012-06-15/146643-36045203.pdf

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

OECD (2006a), Report on the Regulatory Uses and Applications in OECD Member Countries of (Q)SAR Models in the Assessment of New and Existing Chemicals.OECD Environment Health and Safety Publications. Series on Testing and Assessment. No. 58. ENV/JM/MONO(2006)25, OECD, Paris.

OECD (2006b), SIDS Initial Assessment Profile. Long chain alcohols (C6-22 primary aliphatic alcohols, No. 166, Series on Testing and Assessment. ENV/JM/MONO(2012)4/PART5, OECD, Paris.

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

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

OECD (2007a), SIDS Initial Assessment Profile. Alkyl sulfates, alkane sulfonates and α-olefin sulfonates, No. 166, Series on Testing and Assessment. ENV/JM/MONO(2012)4/PART5, OECD, Paris.

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

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

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

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

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

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

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

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

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

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

OECD (2012a), The Adverse Outcome Pathway for Skin Sensitisation Initiated by Covalent Binding to Proteins, OECD Series on Testing and Assessment, No. 168, OECD Publishing, Paris, https://doi.org/10.1787/9789264221444-en.

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

OECD (2012c), SIDS Initial Assessment Profile. Tertiary Amines, http://webnet.oecd.org/HPV/UI/handler.axd?id=ee5efcff-697d-48f7-ab7d-032356a83370

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

OECD (2012e), Report of an OECD workshop on metals specificities in environmental risk assessment

7-8 September 2011, Paris, France Series on Testing & Assessment No. 165. ENV/JM/MONO(2012)2.

OECD (2014a), Guidance Document for Describing Non-Guideline *In Vitro* Test Methods, Series on Testing and Assessment No. 211, ENV/JM/MONO(2014)35, OECD, Paris. Available at http://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm.

OECD (2014b), "SIDS Initial Assessment Profile Soluble Co salts CoCAM 6. http://webnet.oecd.org/hpv/ui/handler.axd?id=e5e60085-1f3f-4df5-92f6-8f32c26c3082

OECD (2014c), Guidance Document on the Validation of (Quantitative) Structure-Activity Relationship [(Q)SAR] Models, OECD Series on Testing and Assessment, No. 69, OECD Publishing, Paris, https://doi.org/10.1787/9789264085442-en.

OECD (2014d), The Guidance Document for Using the OECD (Q)SAR Application Toolbox to Develop Chemical Categories According to the OECD Guidance on Grouping Chemicals, OECD Series on Testing and Assessment, No. 102, OECD Publishing, Paris, https://doi.org/10.1787/9789264221482-en.

OECD (2015). Guidance For Characterising Hydrocarbon Solvents For Assessment Purposes No. 230, OECD Series on Adverse Outcome Pathways No. 1, ENV/JM/MONO(2015)52, OECD, Paris.

OECD (2016a). Users' Handbook supplement to the Guidance Document for developing and accessing Adverse Outcome Pathways, Series on Testing & Assessment No. 233, OECD Series on Adverse Outcome Pathways No. 1, ENV/JM/MONO(2016)12, OECD, Paris.

OECD (2016b). Report on Considerations from Case Studies on Integrated Approaches for Testing and Assessment (IATA), First Review Cycle (2015), Case Studies on Grouping Methods as a Part of IATA, Series on Testing & Assessment No. 250. ENV/JM/MONO(2016)48, OECD, Paris.

OECD (2016c). Approaches on nano grouping/ equivalence/ read-across concepts based on physical-chemical properties (GERA-PC) for regulatory regimes, Series on the Safety of Manufactured Nanomaterials No. 64, ENV/JM/MONO(2016)3, OECD, Paris

OECD (2016d). Grouping and Read-Across for the Hazard Assessment of Manufactured Nanomaterials, Series on the Safety of Manufactured Nanomaterials No. 76, ENV/JM/MONO(2016)59, OECD, Paris

OECD (2016e) Guidance Document on the Reporting of Defined Approaches to be used within IATA. OECD Series on Testing & Assessment No. 255, ENV/JM/MONO(2016)28, OECD, Paris.

OECD (2016f) Case Study on the Use of an Integrated Approach to Testing and Assessment for hepatotoxicity of Allyl Esters. Series on the Safety of Manufactured Nanomaterials No. 76, ENV/JM/MONO(2016)51, OECD, Paris.

OECD (2016g). OECD Guidance For Characterising Hydrocarbon Solvents For Assessment Purposes Series on Testing and Assessment No. 230. ENV/JM/MONO(2015)52, OECD, Paris.

OECD (2017a). Report on Considerations from Case Studies on Integrated Approaches for Testing and Assessment (IATA), Second Review Cycle (2016), Series on Testing & Assessment No. 270. ENV/JM/MONO(2017)22, OECD, Paris.

OECD (2017b). Alternative testing strategies in risk assessment of manufactured nanomaterials: current state of knowledge and research needs to advance their use, Series on the Safety of

Manufactured Nanomaterials No. 80, ENV/JM/MONO(2016)63, OECD, Paris

OECD (2017c). Guidance on Selecting a Strategy for Assessing the Ecological risk of Organometallic and Organic Metal Salt Substances based on their Environmental Fate, Series on Testing & Assessment No. 212, ENV/JM/MONO(2015)2, OECD, Paris

OECD (2017d) Guidance on Grouping of Chemicals. Second Edition, 2017. OECD Series on Testing and Assessment, No. 194, OECD Publishing, Paris, https://doi.org/10.1787/9789264274679-en

OECD (2017e), OECD Guidance for Characterising Oleochemical Substances for Assessment Purposes, OECD Series on Testing and Assessment, No. 193, OECD Publishing, Paris, https://doi.org/10.1787/9789264274778-en.

OECD (2017f), Test No. 318: Dispersion Stability of Nanomaterials in Simulated Environmental Media, OECD Guidelines for the Testing of Chemicals, Section 3, OECD Publishing, Paris, https://doi.org/10.1787/9789264284142-en.

OECD (2017g), Guidance document on developing and assessing adverse outcome pathways, No. 184, Series on Testing and Assessment. ENV/JM/MONO(2013)6, OECD, Paris.

OECD (2018a) Prioritization of Chemicals using the Integrated Approaches for Testing and Assessment (IATA)-based ecological risk classification. Series on Testing & Assessment No. 291. ENV/JM/MONO(2018)27, OECD, Paris.

OECD (2018b), Case study on the use of integrated approaches for testing and assessment (IATA) for estrogenicity of the substituted phenols. Series on Testing and Assessment No. 290. ENV/JM/MONO(2018)26. OECD, Paris.

OECD (2018c) Considerations for Assessing the Risks of Combined Exposure to Multiple Chemicals, Series on Testing and Assessment No. 296, ENV/JM/MONO(2018)37

OECD (2019a), Guiding Principles and Key Elements for Establishing a Weight of Evidence for Chemical Assessment , OECD Series on Testing and Assessment, No. 311, OECD Publishing, Paris, https://doi.org/10.1787/f11597f6-en.

OECD (2019b), Case Study on the Use of an Integrated Approach to Testing and Assessment for Testicular Toxicity of Ethylene Glycol Methyl Ether (EGME)-Related Chemicals, Series on Testing & Assessment No. 308. ENV/JM/MONO(2019)27, OECD, Paris.

OECD (2019c), Guidance Document on Aquatic Toxicity Testing of Difficult Substances and Mixtures, OECD Series on Testing and Assessment, OECD Publishing, Paris, https://doi.org/10.1787/0ed2f88e-en.

OECD (2020a), Overview of Concepts and Available Guidance related to Integrated Approaches to Testing and Assessment (IATA), No. 329, Series on Testing and Assessment. ENV/JM/MONO(2020)25, OECD, Paris.

OECD (2020c), Advancing Adverse Outcome Pathway (AOP) Development for Nanomaterial Risk Assessment and Categorisation Part 1: Final Project Report and Recommendations with Methodology to Prioritise Key Events (KEs) Relevant for Manufactured Nanomaterials, Series on the Safety of Manufactured Nanomaterials No.93, ENV/JM/MONO(2020)33, OECD, Paris

OECD (2020d), Advancing Adverse Outcome Pathway (AOP) Development for Nanomaterial Risk Assessment and Categorisation Part 2: Case Study on Tissue Injury, Series on the Safety of

Manufactured Nanomaterials No.94, ENV/JM/MONO(2020)34, OECD, Paris

OECD (2020e), Advancing Adverse Outcome Pathway (AOP) Development for Nanomaterial Risk Assessment and Categorization Part3: Workshop Report and Recommendations, Series on the Safety of Manufactured Nanomaterials No.95, ENV/JM/MONO(2020)35, OECD, Paris

OECD (2020f), Case Study on the use of Integrated Approaches for Testing and Assessment for Systemic Toxicity Arising from Cosmetic Exposure to Caffeine. Series on Testing & Assessment No 321. ENV/JM/MONO(2020)17, OECD, Paris

OECD (2020g), Case Study on the use of Integrated Approaches for Testing and Assessment to Inform Read-across of p-Alkylphenols: Repeated-Dose Toxicity. Series on Testing & Assessment No 320. ENV/JM/MONO(2020)19, OECD, Paris

OECD (2020h), Case study on the use of integrated approaches for testing and assessment for 90-day rat oral repeated-dose toxicity of chlorobenzene-related chemicals. Series on Testing and Assessment No. 322. ENV/JM/MONO(2020)18, OECD, Paris.

OECD (2020i), Case study on the use of integrated approaches to testing and assessment for prediction of a 90 day repeated dose toxicity study (OECD 408) for 2-ethylbutyric acid using read-across approach from other branched carboxylic acids. Series on Testing and Assessment No. 324. ENV/JM/MONO(2020)20, OECD, Paris.

OECD (2020j), Case study on the use of integrated approaches to testing and assessment for read-across based filling of developmental toxicity data gap for methyl hexanoic acid. Series on Testing and Assessment No. 325. ENV/JM/MONO(2020)21, OECD, Paris.

OECD (2020k), Case study on the Use of Integrated Approaches for Testing and Assessment (IATA) for identification and characterisation of parkinsonian hazard liability of deguelin by an AOP-based testing and read-across approach, OECD Series on Testing and Assessment, No. 326, ENV/JM/MONO(2020)22, OECD, Paris.

OECD (2020l), Case study on the Use of Integrated Approaches for Testing and Assessment (IATA) Mitochondrial Complex-III-mediated neurotoxicity of - Read-Across to other strobilurins, OECD Series on Testing and Assessment, No. 327, ENV/JM/MONO(2020)23, OECD, Paris.

OECD (2020m), Guidance document on aquatic and sediment toxicological testing of nanomaterials, OECD Series on Testing and Assessment, No. 317, ENV/JM/MONO(2020)8, OECD, Paris.

OECD (2020n), Case study on the Use of Integrated Approaches for Testing and Assessment (IATA) for Prediction of a 90 day repeated dose toxicity study (OECD 408) for 2-ethlylbutyric acid using read-across approach from other branched carboxylic acids, OECD Series on Testing and Assessment, No. 324, ENV/JM/MONO(2020)20, OECD, Paris.

OECD (2021a), Guidance document on the characterisation, validation and reporting of Physiologically Based Kinetic (PBK) models for regulatory purposes, OECD Series on Testing and Assessment, No. 331, OECD, Paris, https://doi.org/10.1787/d0de241f-en.

OECD (2021b), Test No. 249: Fish Cell Line Acute Toxicity - The RTgill-W1 cell line assay, OECD Guidelines for the Testing of Chemicals, Section 2, OECD, Paris, https://doi.org/10.1787/c66d5190-en .

OECD (2022), Case study on the Use of Integrated Approaches for Testing and Assessment (IATA) for developmental neurotoxicity hazard characterisation of imidacloprid and the metabolite desnitro-

imidacloprid, OECD Series on Testing and Assessment, No. 366, ENV/CBC/MONO(2022)28, OECD, Paris.

OECD (2023a). OECD Omics Reporting Framework (OORF): Guidance on reporting elements for the regulatory use of omics data from laboratory-based toxicology studies. OECD Environment, Health and Safety Publications Series on Testing & Assessment No. 390, https://doi.org/10.1787/6bb2e6ce-en, OECD, Paris.

OECD (2023b), Supporting Document to the OECD Guideline No. 497 on Defined Approaches for Skin Sensitisation,OECD Series on Testing and Assessment, No. 336, ENV/CBC/MONO(2021)11, OECD, Paris.

OECD (2023c), *Advanced Materials: Working Description*,OECD Series on The Safety of Manufactured Nanomaterials and other Advanced Materials, OECD Publishing, Paris, https://doi.org/10.1787/4b5ba38d-en. No. 104, ENV/CBC/MONO(2022)29, OECD, Paris.

OECD (2024a), *(Q)SAR Assessment Framework: Guidance for the regulatory assessment of (Quantitative) Structure Activity Relationship models and predictions*, OECD Series on Testing and Assessment, No. 386, OECD Publishing, Paris, https://doi.org/10.1787/d96118f6-en

OECD (2024b), *Case Study on the Use of Integrated Approaches for Testing and Assessment (IATA) for Chronic Toxicity and Carcinogenicity of Agrichemicals with Exemplar Case Studies - Ninth Review Cycle (2023)*, OECD Series on Testing and Assessment, No. 402, OECD Publishing, Paris, https://doi.org/10.1787/c3b9ac37-en.

OECD (2025a), *Test No. 442C: In Chemico Skin Sensitisation: Assays addressing the Adverse Outcome Pathway key event on covalent binding to proteins*, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris, https://doi.org/10.1787/9789264229709-en.

OECD (2025b), *Guideline No. 497: Defined Approaches on Skin Sensitisation*, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris, https://doi.org/10.1787/b92879a4-en.

OECD (2025c), *Guidance on Sample Preparation and Dosimetry for Manufactured Nanomaterials, 2025 Edition*, OECD Series on the Safety of Manufactured Nanomaterials and other Advanced Materials, OECD Publishing, Paris, https://doi.org/10.1787/87ec4ecc-en.

Oomen, A. et al. (2015), “Grouping and Read-Across Approaches for Risk Assessment of Nanomaterials”, *International Journal of Environmental Research and Public Health*, Vol. 12/10, pp. 13415-13434, https://doi.org/10.3390/ijerph121013415.

Park, M. et al. (2018), “Development of a systematic method to assess similarity between nanomaterials for human hazard evaluation purposes – lessons learnt”, *Nanotoxicology*, Vol. 12/7, pp. 652-676, https://doi.org/10.1080/17435390.2018.1465142.

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

Patlewicz, G. et al. (2015), “Building scientific confidence in the development and evaluation of read-across”, *Regulatory Toxicology and Pharmacology*, Vol. 72/1, pp. 117-133, https://doi.org/10.1016/j.yrtph.2015.03.015.

Patlewicz, G. et al. (2013a), "Workshop: Use of “read-across” for chemical safety assessment under

REACH", *Regulatory Toxicology and Pharmacology*, Vol. 65/2, pp. 226-228.

Patlewicz, G. et al. (2013b), "Use of category approaches, read-across and (Q)SAR: General considerations", *Regulatory Toxicology and Pharmacology*, Vol. 67/1, pp. 1-12, https://doi.org/10.1016/j.yrtph.2013.06.002.

Patlewicz, G. et al. (2001), "Skin-sensitization structure-activity relationships for aldehydes", *Contact Dermatitis*, Vol. 44/6, pp. 331-336, https://doi.org/10.1034/j.1600-0536.2001.044006331.x.

Patlewicz, G. et al. (2018), "Navigating through the minefield of read-across frameworks: A commentary perspective", *Computational Toxicology*, Vol. 6, pp. 39-54, https://doi.org/10.1016/j.comtox.2018.04.002.

Patlewicz, G. et al. (2007), "TIMES-SS—A promising tool for the assessment of skin sensitization hazard. A characterization with respect to the OECD validation principles for (Q)SARs and an external evaluation for predictivity", *Regulatory Toxicology and Pharmacology*, Vol. 48/2, pp. 225-239, https://doi.org/10.1016/j.yrtph.2007.03.003.

Patlewicz, G. et al. (2017), "Navigating through the minefield of read-across tools: A review of *in silico* tools for grouping", *Computational Toxicology*, Vol. 3, pp. 1-18, https://doi.org/10.1016/j.comtox.2017.05.003.

Patlewicz, G. et al. (2024a), "Development of chemical categories for per- and polyfluoroalkyl substances (PFAS) and the proof-of-concept approach to the identification of potential candidates for tiered toxicological testing and human health assessment", *Computational Toxicology*, Vol. 31, p. 100327, https://doi.org/10.1016/j.comtox.2024.100327.

Patlewicz, G. et al. (2024b), "A systematic analysis of read-across within REACH registration dossiers", *Computational Toxicology*, Vol. 30, p. 100304, https://doi.org/10.1016/j.comtox.2024.100304.

Patlewicz, G. et al. (2010), "Can mutagenicity information be useful in an Integrated Testing Strategy (ITS) for skin sensitization?", *SAR and QSAR in Environmental Research*, Vol. 21/7-8, pp. 619-656, https://doi.org/10.1080/1062936x.2010.528447.

Patlewicz, G. et al. (2023), "Evaluating the utility of a high throughput thiol-containing fluorescent probe to screen for reactivity: A case study with the Tox21 library", *Computational Toxicology*, Vol. 26, p. 100271, https://doi.org/10.1016/j.comtox.2023.100271.

Patlewicz, G. and I. Shah (2023), "Towards systematic read-across using Generalised Read-Across (GenRA)", *Computational Toxicology*, Vol. 25, p. 100258, https://doi.org/10.1016/j.comtox.2022.100258.

Pearce, R. et al. (2017), "httk: *R* Package for High-Throughput Toxicokinetics", *Journal of Statistical Software*, Vol. 79/4, https://doi.org/10.18637/jss.v079.i04.

Peters, A. et al. (2023), "Updating the Chronic Freshwater Ecotoxicity Database and Biotic Ligand Model for Nickel for Regulatory Applications in Europe", *Environmental Toxicology and Chemistry*, Vol. 42/3, pp. 566-580, https://doi.org/10.1002/etc.5561.

Pham, L. et al. (2020), "Variability in *in vivo* studies: Defining the upper limit of performance for predictions of systemic effect levels", *Computational Toxicology*, Vol. 15, p. 100126, https://doi.org/10.1016/j.comtox.2020.100126.

Pradeep, P. et al. (2019), "Integrating data gap filling techniques: A case study predicting TEFs for

neurotoxicity TEQs to facilitate the hazard assessment of polychlorinated biphenyls", *Regulatory Toxicology and Pharmacology*, Vol. 101, pp. 12-23, https://doi.org/10.1016/j.yrtph.2018.10.013.

Rager, J. and C. Rider (2023), "Wrangling whole mixtures risk assessment: Recent advances in determining sufficient similarity", *Current Opinion in Toxicology*, Vol. 35, p. 100417, https://doi.org/10.1016/j.cotox.2023.100417.

Rasmussen, K. et al. (1999), "Compilation of EINECS: Descriptions and definitions used for UVCB substances: Complex reaction products, plant products, (post-reacted) naturally occurring substances, micro-organisms, petroleum products, soaps and detergents, and metallic compounds", *Toxicological & Environmental Chemistry*, Vol. 69/3-4, pp. 403-416, https://doi.org/10.1080/02772249909358721.

Rasmussen, K. et al. (2019), "Developing OECD test guidelines for regulatory testing of nanomaterials to ensure mutual acceptance of test data", *Regulatory Toxicology and Pharmacology*, Vol. 104, pp. 74-83, https://doi.org/10.1016/j.yrtph.2019.02.008.

Rasmussen, K., J. Riego Sintes and H. Rauscher (2024), "How nanoparticles are counted in global regulatory nanomaterial definitions", *Nature Nanotechnology*, Vol. 19/2, pp. 132-138, https://doi.org/10.1038/s41565-023-01578-x.

Rathman, J., C. Yang and H. Zhou (2018), "Dempster-Shafer theory for combining *in silico* evidence and estimating uncertainty in chemical risk assessment", *Computational Toxicology*, Vol. 6, pp. 16-31, https://doi.org/10.1016/j.comtox.2018.03.001.

RCC (2013a), Nanotechnology Initiative Work Element 2, Priority Setting: Development of a Joint Nanomaterials Classification Scheme. Final report, Regulatory Cooperation Council (RCC).

RCC (2013b), Nanotechnology Initiative Work Element 3: Risk Assessment/Risk Management. Final report, Regulatory Cooperation Council (RCC). http://doi.org/10.1007/s11051-012-1029-8.

RCC (2014), Work Element 2, Priority Setting: Development of a Joint Nanomaterials Classification Scheme. https://www.ic.gc.ca/eic/site/063.nsf/vwapj/nano2_e.pdf/$file/nano2_e.pdf

Richard, A. et al. (2022), "Identification of Branched and Linear Forms of PFOA and Potential Precursors: A User-Friendly SMILES Structure-based Approach", *Frontiers in Environmental Science*, Vol. 10, https://doi.org/10.3389/fenvs.2022.865488.

Ridder, L. and M. Wagener (2008), "SyGMa: Combining Expert Knowledge and Empirical Scoring in the Prediction of Metabolites", *ChemMedChem*, Vol. 3/5, pp. 821-832, https://doi.org/10.1002/cmdc.200700312.

Robinson, B. (1999), "Soil Amendments Affecting Nickel and Cobalt Uptake by Berkheya coddii: Potential Use for Phytomining and Phytoremediation", *Annals of Botany*, Vol. 84/6, pp. 689-694, https://doi.org/10.1006/anbo.1999.0970.

Roe, H. et al. (2024), *A Systematic Analysis of Read-Across Adaptations in Testing Proposal Evaluations by the European Chemicals Agency*, Cold Spring Harbor Laboratory, https://doi.org/10.1101/2024.08.29.610278.

Roell, K. et al. (2022), "Development of the InTelligence And Machine LEarning (TAME) Toolkit for Introductory Data Science, Chemical-Biological Analyses, Predictive Modeling, and Database Mining for Environmental Health Research", *Frontiers in Toxicology*, Vol. 4, https://doi.org/10.3389/ftox.2022.893924.

Rogers, D. and M. Hahn (2010), "Extended-Connectivity Fingerprints", *Journal of Chemical Information and Modeling*, Vol. 50/5, pp. 742-754, https://doi.org/10.1021/ci100050t.

Rovida, C. (2021), "NAM-supported read-across: From case studies to regulatory guidance in safety assessment", *ALTEX*, pp. 140-150, https://doi.org/10.14573/altex.2010062.

Rovida, C. (2020), "Internationalization of read-across as a validated new approach method (NAM) for regulatory toxicology", *ALTEX*, https://doi.org/10.14573/altex.1912181.

Rowan-Carroll, A. et al. (2021), "High-Throughput Transcriptomic Analysis of Human Primary Hepatocyte Spheroids Exposed to Per- and Polyfluoroalkyl Substances as a Platform for Relative Potency Characterization", *Toxicological Sciences*, Vol. 181/2, pp. 199-214, https://doi.org/10.1093/toxsci/kfab039.

Ruggiero, E. et al. (2022), "Food contact of paper and plastic products containing SiO_2, Cu-Phthalocyanine, Fe_2O_3, $CaCO_3$: Ranking factors that control the similarity of form and rate of release", *NanoImpact*, Vol. 25, p. 100372, https://doi.org/10.1016/j.impact.2021.100372.

Russom, C. et al. (1997), "Predicting modes of toxic action from chemical structure: Acute toxicity in the fathead minnow (*Pimephales promelas*)", *Environmental Toxicology and Chemistry*, Vol. 16/5, pp. 948-967, https://doi.org/10.1002/etc.5620160514.

Ryan, K. et al. (2019), "Evaluating Sufficient Similarity of Botanical Dietary Supplements: Combining Chemical and *In Vitro* Biological Data", *Toxicological Sciences*, Vol. 172/2, pp. 316-329, https://doi.org/10.1093/toxsci/kfz189.

Ryan, N. et al. (2016), "Moving Toward Integrating Gene Expression Profiling Into High-Throughput Testing: A Gene Expression Biomarker Accurately Predicts Estrogen Receptor α Modulation in a Microarray Compendium", *Toxicological Sciences*, Vol. 151/1, pp. 88-103, https://doi.org/10.1093/toxsci/kfw026.

Safe, S. (1991), "Polychlorinated dibenzo-*p*-dioxins and related compounds: Sources, environmental distribution and risk assessment", *Journal of Environmental Science and Health, Part C*, Vol. 9/2, pp. 261-302, https://doi.org/10.1080/10590509109373409.

Saghir, S. et al. (2012), "Assessment of diurnal systemic dose of agrochemicals in regulatory toxicity testing – An integrated approach without additional animal use", *Regulatory Toxicology and Pharmacology*, Vol. 63/2, pp. 321-332, https://doi.org/10.1016/j.yrtph.2012.03.004.

Saliner, G.et al., (2009), Review of computational approaches for predicting the physicochemical and biological properties of nanoparticles, EUR 23974 EN. https://publications.jrc.ec.europa.eu/repository/handle/JRC52120

Salminen, R. et al., (2005), Geochemical atlas of Europe. Part 1—background information, methodology and maps. Geological Survey of Finland. Otamedia Oy, Espoo 2005, 525. http://weppi.gtk.fi/publ/foregsatlas/articles.php

Salvito, D. (2007), Approaches to Chemical Categorization: An Illustrative Example of Approaches Used by the Fragrance Industry. In: A Compendium of Case Studies that helped to shape the REACH Guidance on Chemical Categories and Read Across. Worth A. & Patlewicz G. (Eds.), EUR report no 22481 EN.

Salvito, D. et al. (2020), "Improving the Environmental Risk Assessment of Substances of Unknown or Variable Composition, Complex Reaction Products, or Biological Materials", *Environmental*

Toxicology and Chemistry, Vol. 39/11, pp. 2097-2108, https://doi.org/10.1002/etc.4846.

Salvito, D. et al. (2022), "The Path to UVCB Ecological Risk Assessment: Grappling with Substance Characterization", *Environmental Toxicology and Chemistry*, Vol. 41/11, pp. 2649-2657, https://doi.org/10.1002/etc.5462.

Sapounidou, M. et al. (2021), "Development of an Enhanced Mechanistically Driven Mode of Action Classification Scheme for Adverse Effects on Environmental Species", *Environmental Science & Technology*, Vol. 55/3, pp. 1897-1907, https://doi.org/10.1021/acs.est.0c06551.

Schoeters, I. and V. Verougstraete (2007), Case studies on inorganic metal compounds, A Compendium of Case Studies that helped to shape the REACH Guidance on Chemical Categories and Read Across. Worth A & Patlewicz G (Eds), EUR report no 22481 EN.

Schultz, T. (2010), "Adverse Outcome Pathways: A Way of Linking Chemical Structure to *In Vivo* Toxicological Hazards", in *In Silico* Toxicology, The Royal Society of Chemistry, https://doi.org/10.1039/9781849732093-00346.

Schultz, T. et al. (2015), "A strategy for structuring and reporting a read-across prediction of toxicity", *Regulatory Toxicology and Pharmacology*, Vol. 72/3, pp. 586-601, https://doi.org/10.1016/j.yrtph.2015.05.016.

Schultz, T. et al. (2002), "Structure−Toxicity Relationships for Aliphatic Chemicals Evaluated with *Tetrahymena pyriformis*", *Chemical Research in Toxicology*, Vol. 15/12, pp. 1602-1609, https://doi.org/10.1021/tx025589p.

Schultz, T. et al. (2005), "Structure−Toxicity Relationships for the Effects to *Tetrahymena pyriformis* of Aliphatic, Carbonyl-Containing, α,β-Unsaturated Chemicals", *Chemical Research in Toxicology*, Vol. 18/2, pp. 330-341, https://doi.org/10.1021/tx049833j.

Schultz, T., A. Richarz and M. Cronin (2019), "Assessing uncertainty in read-across: Questions to evaluate toxicity predictions based on knowledge gained from case studies", *Computational Toxicology*, Vol. 9, pp. 1-11, https://doi.org/10.1016/j.comtox.2018.10.003.

Schwirn, K. and D. Völker (2019), "Hazard grouping of nanomaterials for daphnia and algae toxicity: lessons learned from scientific projects for regulatory applicability", *Environmental Sciences Europe*, Vol. 31/1, https://doi.org/10.1186/s12302-019-0226-9.

Schymanski, E. et al. (2015), "Non-target screening with high-resolution mass spectrometry: critical review using a collaborative trial on water analysis", *Analytical and Bioanalytical Chemistry*, Vol. 407/21, pp. 6237-6255, https://doi.org/10.1007/s00216-015-8681-7.

Sedykh, A. et al. (2011), "Use of *in Vitro* HTS-Derived Concentration–Response Data as Biological Descriptors Improves the Accuracy of QSAR Models of *in Vivo* Toxicity", *Environmental Health Perspectives*, Vol. 119/3, pp. 364-370, https://doi.org/10.1289/ehp.1002476.

Seleci, D. et al. (2022), "Determining nanoform similarity via assessment of surface reactivity by abiotic and *in vitro* assays", NanoImpact, Vol. 26, p. 100390, https://doi.org/10.1016/j.impact.2022.100390.

Serafimova, R. et al. (2007), "Identification of the Structural Requirements for Mutagencitiy, by Incorporating Molecular Flexibility and Metabolic Activation of Chemicals. II. General Ames Mutagenicity Model", *Chemical Research in Toxicology*, Vol. 20/4, pp. 662-676, https://doi.org/10.1021/tx6003369.

Shafer, G. (1976), "A Mathematical Theory of Evidence", Princeton University Press.

Shah, I. et al. (2016), “Systematically evaluating read-across prediction and performance using a local validity approach characterized by chemical structure and bioactivity information”, *Regulatory Toxicology and Pharmacology*, Vol. 79, pp. 12-24, https://doi.org/10.1016/j.yrtph.2016.05.008.

Shaw, S. et al. (2008), “Perturbational profiling of nanomaterial biologic activity”, *Proceedings of the National Academy of Sciences*, Vol. 105/21, pp. 7387-7392, https://doi.org/10.1073/pnas.0802878105.

Simmons, J. (2012), “Toxicity of major cations and anions (Na^+, K^+, Ca^{2+}, Cl^-, and SO_4^{2-}) to a macrophyte and an alga”, *Environmental Toxicology and Chemistry*, Vol. 31/6, pp. 1370-1374, https://doi.org/10.1002/etc.1819.

Simon, T., J. Britt and R. James (2007), “Development of a neurotoxic equivalence scheme of relative potency for assessing the risk of PCB mixtures”, *Regulatory Toxicology and Pharmacology*, Vol. 48/2, pp. 148-170, https://doi.org/10.1016/j.yrtph.2007.03.005.

Smeltz, M., J. Wambaugh and B. Wetmore (2023), “Plasma Protein Binding Evaluations of Per- and Polyfluoroalkyl Substances for Category-Based Toxicokinetic Assessment”, *Chemical Research in Toxicology*, Vol. 36/6, pp. 870-881, https://doi.org/10.1021/acs.chemrestox.3c00003.

Smith, S. (2009), “Chiral Toxicology: It’s the Same Thing…Only Different”, *Toxicological Sciences*, Vol. 110/1, pp. 4-30, https://doi.org/10.1093/toxsci/kfp097.

Sobus, J. et al. (2017), “Integrating tools for non-targeted analysis research and chemical safety evaluations at the US EPA”, *Journal of Exposure Science & Environmental Epidemiology*, Vol. 28/5, pp. 411-426, https://doi.org/10.1038/s41370-017-0012-y.

Song, Y. et al. (2022), “Similarity assessment of metallic nanoparticles within a risk assessment framework: A case study on metallic nanoparticles and lettuce”, *NanoImpact*, Vol. 26, p. 100397, https://doi.org/10.1016/j.impact.2022.100397.

Sperber, S. et al. (2019), “Metabolomics as read-across tool: An example with 3-aminopropanol and 2-aminoethanol”, *Regulatory Toxicology and Pharmacology*, Vol. 108, p. 104442, https://doi.org/10.1016/j.yrtph.2019.104442.

Spurgeon, D., E. Lahive and C. Schultz (2020), “Nanomaterial Transformations in the Environment: Effects of Changing Exposure Forms on Bioaccumulation and Toxicity”, *Small*, Vol. 16/36, https://doi.org/10.1002/smll.202000618.

Stone, V. et al. (2020), “A framework for grouping and read-across of nanomaterials- supporting innovation and risk assessment”, *Nano Today*, Vol. 35, p. 100941, https://doi.org/10.1016/j.nantod.2020.100941.

Stone, V., H. Johnston and R. Schins (2009), “Development of *in vitro* systems for nanotoxicology: methodological considerations”, *Critical Reviews in Toxicology*, Vol. 39/7, pp. 613-626, https://doi.org/10.1080/10408440903120975.

Stone, V. et al. (2014), “ITS-NANO - Prioritising nanosafety research to develop a stakeholder driven intelligent testing strategy”, *Particle and Fibre Toxicology*, Vol. 11/1, p. 9, https://doi.org/10.1186/1743-8977-11-9.

Tate, T. et al. (2021), “Repeat-dose toxicity prediction with Generalized Read-Across (GenRA) using

targeted transcriptomic data: A proof-of-concept case study", *Computational Toxicology*, Vol. 19, p. 100171, https://doi.org/10.1016/j.comtox.2021.100171.

Todeschini, R. et al. (2012), "Similarity Coefficients for Binary Chemoinformatics Data: Overview and Extended Comparison Using Simulated and Real Data Sets", *Journal of Chemical Information and Modeling*, Vol. 52/11, pp. 2884-2901, https://doi.org/10.1021/ci300261r.

Trimmer, G.W. (2008), "Whole-Body Inhalation Developmental Toxicity Study in Rats with Baseline Gasoline Vapor Condensate (MRD-00-695)", ExxonMobil Biomedical Sciences, Inc. Study No 169534; sponsored by the American Petroleum Institute

Tsakovska, I. and A. Worth (2007), Review of examples of read-across/category approaches discussed by the TC/NES. In: A Compendium of Case Studies that helped to shape the inorganic metal REACH Guidance on Chemical Categories and Read Across, Worth A. & Patlewicz G. (Eds.), EUR report no 22481 EN.

United Nations (2023), Globally Harmonised System of Classification and Labelling of Chemicals (GHS), Tenth revised version, 2023; GHS/Rev.10e (https://unece.org/).

Tsiliki, G. et al. (2022), "Bayesian based similarity assessment of nanomaterials to inform grouping", *NanoImpact*, Vol. 25, p. 100389, https://doi.org/10.1016/j.impact.2022.100389.

van den Brule, S. et al. (2022), "A tiered approach to investigate the inhalation toxicity of cobalt substances. Tier 2a: Grouping cobalt compounds based on their capacity to stabilize HIF-1α in human alveolar epithelial cells *in vitro*", *Regulatory Toxicology and Pharmacology*, Vol. 130, p. 105121, https://doi.org/10.1016/j.yrtph.2022.105121.

van der Stel, W. (2021), "New approach methods supporting read-across: Two neurotoxicity AOP-based IATA case studies", ALTEX, https://doi.org/10.14573/altex.2103051.

van der Weiden, M. et al. (1992), "Concurrence of P450 IAl induction and toxic effects after administration of a low dose of 2,3,7,8-tetrachlorodibenzo-P-dioxin (TCDD) in the rainbow trout (Oncorhynchus mykiss)", *Aquatic Toxicology*, Vol. 24/1-2, pp. 123-142, https://doi.org/10.1016/0166-445x(92)90020-n.

van Mourik, L. et al. (2016), "Chlorinated paraffins in the environment: A review on their production, fate, levels and trends between 2010 and 2015", *Chemosphere*, Vol. 155, pp. 415-428, https://doi.org/10.1016/j.chemosphere.2016.04.037.

van Ravenzwaay, B. et al. (2016), "Metabolomics as read-across tool: A case study with phenoxy herbicides", *Regulatory Toxicology and Pharmacology*, Vol. 81, pp. 288-304, https://doi.org/10.1016/j.yrtph.2016.09.013.

Verdonck, F. et al. (2017), "MeClas: An online tool for hazard identification and classification of complex inorganic metal-containing materials", *Regulatory Toxicology and Pharmacology*, Vol. 89, pp. 232-239, https://doi.org/10.1016/j.yrtph.2017.07.015.

Verdon, R. et al. (2022), "The application of existing genotoxicity methodologies for grouping of nanomaterials: towards an integrated approach to testing and assessment", *Particle and Fibre Toxicology*, Vol. 19/1, https://doi.org/10.1186/s12989-022-00476-9.

Verhaar, H., C. van Leeuwen and J. Hermens (1992), "Classifying environmental pollutants", *Chemosphere*, Vol. 25/4, pp. 471-491, https://doi.org/10.1016/0045-6535(92)90280-5.

Verougstraete, V. et al. (2022), “A tiered approach to investigate the inhalation toxicity of cobalt substances. Tier 1: Bioaccessibility testing”, *Regulatory Toxicology and Pharmacology*, Vol. 129, p. 105124, https://doi.org/10.1016/j.yrtph.2022.105124.

Viant, M. et al. (2024a), “Utilizing Omics Data for Chemical Grouping”, *Environmental Toxicology and Chemistry*, Vol. 43/10, pp. 2094-2104, https://doi.org/10.1002/etc.5959.

Viant, M. et al. (2024b), “Demonstrating the reliability of *in vivo* metabolomics based chemical grouping: towards best practice”, *Archives of Toxicology*, Vol. 98/4, pp. 1111-1123, https://doi.org/10.1007/s00204-024-03680-y.

Viegas, V. et al. (2022), “A tiered approach to investigate the inhalation toxicity of cobalt substances. Tier 3: Inflammatory response following acute inhalation exposure correlates with lower tier data”, *Regulatory Toxicology and Pharmacology*, Vol. 130, p. 105127, https://doi.org/10.1016/j.yrtph.2022.105127.

Visser, M. et al. (2022), “Towards health-based nano reference values (HNRVs) for occupational exposure: Recommendations from an expert panel”, *NanoImpact*, Vol. 26, p. 100396, https://doi.org/10.1016/j.impact.2022.100396.

Vrijenhoek, N. (2022), “Application of high-throughput transcriptomics for mechanism-based biological read-across of short-chain carboxylic acid analogues of valproic acid”, *ALTEX*, https://doi.org/10.14573/altex.2107261.

Walker, M. and R. Peterson (1991), “Potencies of polychlorinated dibenzo-p-dioxin, dibenzofuran, and biphenyl congeners, relative to 2,3,7,8-tetrachlorodibenzo-p-dioxin, for producing early life stage mortality in rainbow trout (*Oncorhynchus mykiss*)”, *Aquatic Toxicology*, Vol. 21/3-4, pp. 219-237, https://doi.org/10.1016/0166-445x(91)90074-j.

Wang, C. et al. (2015), “An invisible soil acidification: Critical role of soil carbonate and its impact on heavy metal bioavailability”, *Scientific Reports*, Vol. 5/1, https://doi.org/10.1038/srep12735.

Wang, N. et al. (2012), “Application of computational toxicological approaches in human health risk assessment. I. A tiered surrogate approach”, *Regulatory Toxicology and Pharmacology*, Vol. 63/1, pp. 10-19, https://doi.org/10.1016/j.yrtph.2012.02.006.

Warheit, D. et al. (2007), “Pulmonary toxicity study in rats with three forms of ultrafine-TiO2 particles: Differential responses related to surface properties”, *Toxicology*, Vol. 230/1, pp. 90-104, https://doi.org/10.1016/j.tox.2006.11.002.

Webster, F. et al. (2019), “Predicting estrogen receptor activation by a group of substituted phenols: An integrated approach to testing and assessment case study”, *Regulatory Toxicology and Pharmacology*, Vol. 106, pp. 278-291, https://doi.org/10.1016/j.yrtph.2019.05.017.

Wheeler, M. (2018), “Bayesian Additive Adaptive Basis Tensor Product Models for Modeling High Dimensional Surfaces: An Application to High-Throughput Toxicity Testing”, *Biometrics*, Vol. 75/1, pp. 193-201, https://doi.org/10.1111/biom.12942.

WHO (2018), "Guidance document on evaluating and expressing uncertainty in hazard characterization, 2nd edition". https://apps.who.int/iris/handle/10665/259858

Wick, P. et al. (2007), “The degree and kind of agglomeration affect carbon nanotube cytotoxicity”, *Toxicology Letters*, Vol. 168/2, pp. 121-131, https://doi.org/10.1016/j.toxlet.2006.08.019.

Williams, A. et al. (2017), “The CompTox Chemistry Dashboard: a community data resource for environmental chemistry”, *Journal of Cheminformatics*, Vol. 9/1, https://doi.org/10.1186/s13321-017-0247-6.

Williams, R. et al. (2019), “Models for assessing engineered nanomaterial fate and behaviour in the aquatic environment”, *Current Opinion in Environmental Sustainability*, Vol. 36, pp. 105-115, https://doi.org/10.1016/j.cosust.2018.11.002.

Wohlleben, W. et al. (2019), “The nanoGRAVUR framework to group (nano)materials for their occupational, consumer, environmental risks based on a harmonized set of material properties, applied to 34 case studies”, *Nanoscale*, Vol. 11/38, pp. 17637-17654, https://doi.org/10.1039/c9nr03306h.

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

Worth, A. et al. (2017), "Evaluation of the availability and applicability of computational approaches in the safety assessment of nanomaterials". JRC Technical Reports, EUR 28617 EN, https://data.europa.eu/doi/10.2760/248139

Wu, S. et al. (2010), “A framework for using structural, reactivity, metabolic and physicochemical similarity to evaluate the suitability of analogs for SAR-based toxicological assessments”, *Regulatory Toxicology and Pharmacology*, Vol. 56/1, pp. 67-81, https://doi.org/10.1016/j.yrtph.2009.09.006.

Wu, S. et al. (2022), “Structure-activity relationship read-across and transcriptomics for branched carboxylic acids”, *Toxicological Sciences*, Vol. 191/2, pp. 343-356, https://doi.org/10.1093/toxsci/kfac139.

Wu, S. et al. (2013), “Framework for Identifying Chemicals with Structural Features Associated with the Potential to Act as Developmental or Reproductive Toxicants”, *Chemical Research in Toxicology*, Vol. 26/12, pp. 1840-1861, https://doi.org/10.1021/tx400226u.

Yang, C. et al. (2021), “COSMOS next generation – A public knowledge base leveraging chemical and biological data to support the regulatory assessment of chemicals”, *Computational Toxicology*, Vol. 19, p. 100175, https://doi.org/10.1016/j.comtox.2021.100175.

Yang, C. et al. (2023), “High Throughput Read-Across for Screening a Large Inventory of Related Structures by Balancing Artificial Intelligence/Machine Learning and Human Knowledge”, *Chemical Research in Toxicology*, Vol. 36/7, pp. 1081-1106, https://doi.org/10.1021/acs.chemrestox.3c00062.

Yang, C. et al. (2015), “New Publicly Available Chemical Query Language, CSRML, To Support Chemotype Representations for Application to Data Mining and Modeling”, *Journal of Chemical Information and Modeling*, Vol. 55/3, pp. 510-528, https://doi.org/10.1021/ci500667v.

Yap, C. (2010), “PaDEL-descriptor: An open source software to calculate molecular descriptors and fingerprints”, *Journal of Computational Chemistry*, Vol. 32/7, pp. 1466-1474, https://doi.org/10.1002/jcc.21707.

Yauk, C. et al. (2016), “Application of the TGx-28.65 transcriptomic biomarker to classify genotoxic and non-genotoxic chemicals in human TK6 cells in the presence of rat liver S9”, *Environmental and Molecular Mutagenesis*, Vol. 57/4, pp. 243-260, https://doi.org/10.1002/em.22004.

Yordanova, D. et al. (2019), “Using metabolic information for categorization and read-across in the

OECD QSAR Toolbox", *Computational Toxicology*, Vol. 12, p. 100102, https://doi.org/10.1016/j.comtox.2019.100102.

Yordanova, D. et al. (2021), "Assessing metabolic similarity for read-across predictions", *Computational Toxicology*, Vol. 18, p. 100160, https://doi.org/10.1016/j.comtox.2021.100160.

Zabel, E., P. Cook and R. Peterson (1995), "Toxic equivalency factors of polychlorinated dibenzo-p-dioxin, dibenzofuran and biphenyl congeners based on early life stage mortality in rainbow trout (Oncorhynchus mykiss)", *Aquatic Toxicology*, Vol. 31/4, pp. 315-328, https://doi.org/10.1016/0166-445x(94)00075-2.

Zabeo, A. et al. (2022), "Ordered weighted average based grouping of nanomaterials with Arsinh and dose response similarity models", *NanoImpact*, Vol. 25, p. 100370, https://doi.org/10.1016/j.impact.2021.100370.

Zhu, H. (2016), "Supporting read-across using biological data", *ALTEX*, pp. 167-182, https://doi.org/10.14573/altex.1601252.

Zhu, H. et al. (2008), "Use of Cell Viability Assay Data Improves the Prediction Accuracy of Conventional Quantitative Structure–Activity Relationship Models of Animal Carcinogenicity", *Environmental Health Perspectives*, Vol. 116/4, pp. 506-513, https://doi.org/10.1289/ehp.10573.

Appendix

Computational tools

Table 31. Computational tools supporting the development of the categories – related to specific aspects of different endpoints[139]

Endpoint	Approaches and tools
Physicochemical parameters	Physicochemical parameters play a critical role in addressing many aspects of the substance's behaviour and in characterising the chemical similarity for read-across purposes. Basic physicochemical properties provide key information for the assessment of a chemical and in particular for the assessment of the environmental properties. Consequently, experimental data or valid (Q)SAR predictions should normally be available (or should be reasonably obtainable). However, there may occasionally be practical problems, especially for UVCBs, when the use of read-across techniques will be required. Vapour pressure, *log* K_{ow}, water solubility, MW, *pKa* are critical when considering bioaccumulation in the environment and absorption in the animal/human organism and should be addressed for the category members. Tools: EPI Suite, CompTox Dashboard, ACD/ Percepta, ADMET Predictor, QSAR Toolbox, ChemAxon, ChemProp, T.E.S.T., OPERA, VEGA, CORINA Symphony Community Edition, iSafeRat, ProtoPred.
Aquatic toxicity	To facilitate and justify the read-across approach for the aquatic endpoint toxicity, tools like rule base schemes, (Q)SAR-based and WoE approaches are helpful to indicate the mode of action of the substance, and elements that can be used to demonstrate similarity between two or more analogues. A combination of (Q)SAR model and empirical data might also be considered to test any hypothesis and strengthen the overall strategy. Tools: Catalogic, TIMES, ECOSAR, Aquatic toxicity classification by ECOSAR, CompTox Dashboard, iSafeRat, Biovia Discovery Studio (TOPKAT), T.E.S.T., VEGA, ACD/Percepta, Instem (formerly Leadscope), ProtoPredTIMES, Verhaar rulebase within Toxtree or as a profiler within the QSAR Toolbox, OASIS Acute Aquatic Toxicity Mode of Action Profiler within the QSAR Toolbox, OPERA (CERAPP and CoMPARA), LAZAR Danish (Q)SAR Database and Models free websites with predictions and models for acute toxicity to *Fathead minnow*, *Daphnia magna*, and *Pseudokirchneriella subcapitata*.

[139] ECETOC TR116 is acknowledged as a significant source of information for this table. https://www.ecetoc.org/wp-content/uploads/2021/10/ECETOC-TR-116-Category-approaches-Read-across-QSAR.pdf

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

Endpoint	Approaches and tools
	Tools: CompTox Dashboard, T.E.S.T., Biovia Discovery Studio (TOPKAT), QSAR Toolbox, ChemTunes.ToxGPS, OPERA, TIMES (TIssue MEtabolism Simulator), ACD/Percepta, Instem (formerly Leadscope).
Acute **Inhalation route**	The physicochemical properties of the substance and chemical reactivity are major determinants of toxicity. Particle size, vapour pressure, and water solubility are all important especially in the case of volatile substances, both solid aerosol and liquid aerosols may be respirable and trigger hazard. A read-across approach of volatile substances could consider information from other endpoints where narcosis and electrophilic reactivity play a role. In this case, (QSAR) models can be used to demonstrate whether a general narcosis mode of action was relevant. For non-volatiles substances, the read-across approach can be supported by information from other routes of exposure, i.e. dermal and oral and there are default models available. Tools: Biovia Discovery Studio (TOPKAT), QSAR Toolbox.
Irritation **Skin**	Physicochemical information is important for evaluation of the endpoint - especially information on pH, where low and high values are sufficient to determine a substances likely skin and eye irritant / corrosive potential, e.g. pH <2 or >11.5 are considered as corrosive. In order to predict the absence of skin corrosion/irritation, the read-across approach can be supported using SARs models, together with the physicochemical properties of the substance. Danish (Q)SAR Database and Models free websites with predictions and model for severe versus mild/no skin irritation. Tools: Derek Nexus, iSafeRat, ACD/Percepta, Instem (formerly Leadscope), Eye and Skin irritation inclusion and exclusion rules by BfR Profiler with the QSAR Toolbox
Irritation **Eye**	The same considerations as for skin irritancy and physicochemical parameters apply. Eye corrosion/irritation can be demonstrated by using SARs models and the physicochemical properties of the substance. Alternative *In vitro* methods are available to support read-across. Tools: Biovia Discovery Studio (TOPKAT), BfR rulebase within Toxtree, Eye and Skin irritation inclusion and exclusion rules by BfR Profiler with the QSAR Toolbox, Derek Nexus, ACD/Percepta, Instem (formerly Leadscope), iSafeRat, IMES (TIssue MEtabolism Simulator)
Sensitisation **Skin**	A WoE strategy can be applied using *in vivo* and *in vitro* data and evaluating the physicochemical profile of the substance. In 2013, EURL ECVAM published its Strategy for Replacement of Animal Testing for Skin Sensitisation Hazard Identification and Classification (Casati et al., 2013). Databases, (Q)SAR models may be used to identify additional substances to use in the read-across approach based on existing data and similarity in physicochemical structure. (Q)SAR models and *in vitro* methods are focused mainly on reactivity = Protein binding, metabolic activation, internalisation and processing by Langerhans cells (LC), transport of antigen by LC to lymph node and activation of T lymphocyte.

Endpoint	Approaches and tools
	Tools: CAESAR Model for Skin Sensitisation, TIMES, Biovia Discovery Studio (TOPKAT), Derek Nexus, MCASE, SMARTS alerts within Toxtree, Protein binding profilers within the QSAR Toolbox, ChemTunes.ToxGPS, Instem (formerly Leadscope), VEGA, iSafeRat Toolbox: Protein binding by OECD profiler, Protein binding by OASIS profiler and Protein binding alerts for skin sensitisation by OASIS. Danish (Q)SAR Database contains predictions from the commercial MultiCASE model A33, remodelled by DTU in Instem (formerly Leadscope) and SciQSAR in agreement with MultiCASE. Furthermore, the Danish (Q)SAR Database and Models websites contains predictions/models for a number of models for skin sensitisation CLP (GHS) classifications developed by DTU for DK-EPA.
Mutagenicity/Genotoxicity	Data from *in vitro* tests are usually available or should be acquired for strategic category members. A large body of experimental data on mutagenicity has allowed identification of structural alerts for mutagenicity and can be considered within a category as part of the supporting evidence. (Q)SAR models are able to make prediction of mutagenicity testing (*Salmonella*) and *in vivo* mutagenicity and may provide supporting evidence. Tools: CAESAR, TIMES, Biovia Discovery Studio (TOPKAT), Derek Nexus, Sarah Nexus, Acrostic, META Ultra, MCASE, T.E.S.T., LAZAR, ChemProp, Leadscope Model Applier (Instem), Benigni/Bossa Rulebase within Toxtree, ToxMIC-ISS plug-in allows the identification of Structure Alerts for the *in vivo* micronucleus assay within Toxtree Profilers within the QSAR Toolbox, ChemTunes.ToxGPS Toolbox: DNA binding by OASIS profiler and DNA Alerts for Ames, CA and MN by OASIS, DNA Binding by OECD profiler, i*n vitro*, *in vivo* and Carcinogenicity/mutagenicity alerts by ISS. Danish (Q)SAR Database and Models free websites (a range of genotox endpoints both *in vivo* and *in vitro*, are covered, in the database also carcinogenicity in rat/mouse in male/female).
Repeated Dose Toxicity	The source chemical study(ies) need to be reviewed for fit for purpose for the read-across considering, test material, route of exposure, test species, study type and validity, protocol, extent of observations, in order to determine if a study is a suitable source for read-across within a category. The similarity in the toxicological profile across all the human health endpoints will be considered as well as ADME information on informing on a suitability of the read-across to target chemicals. At the moment, there are no *in vitro* methods available that can replace *in vivo* repeated dose toxicity data. However, the SEURAT-1 Research Initiative of the EU funded within the 7th Framework Programme is currently conducting six complementary research projects aimed at ultimately replacing animals in repeated dose toxicity testing. (Q)SAR models might be used to demonstrate similarity in reactivity and support existing data. Tools: Biovia Discovery Studio (TOPKAT) (LOAEL, MDT), Derek Nexus, ChemProp, ChemTunes.ToxGPS, LAZAR, HESS Profiler within the QSAR Toolbox, HESS, Fraunhofer database.

Endpoint	Approaches and tools
Reproductive and developmental Toxicity	In building a hypothesis on reproductive toxicity, structural, functional as well as ADME considerations will be used. Gross pathology and histopathology data from repeated dose toxicity studies should also be used. As reproductive and developmental toxicity are complex endpoints and the mode of action is not usually known, currently read-across must be based upon experimental data. Screening studies (OECD TG 421 and TG 422) on category members can provide useful data and provide confidence in the read-across of higher tier studies from source chemicals and test the validity any read-across hypothesis for the endpoint. (Q)SAR models have been developed and may be useful for supporting trends seen in existing data for the category members. Developmental and Reproductive Toxicity (DART) Decision Tree profilers (Wu et al., 2013), and COSMOS NG Database are additional tools to support the Read-across. Currently *in vitro* assays only can be used support specific outcomes on reproductive organs. However, a “Feasibility Study" which concluded the ReProTect an Integrated Project of the EU (funded within the 6th Framework Programme) identified a test battery of 14 *in vitro* assays that allowed a robust prediction of adverse effects on fertility and embryonic development of 10 test chemicals with toxicologically well-documented profiles *in vivo*. Tools: ADMET Predictor, ACD Percepta, ChemProp, CompTox Dashboard, Derek Nexus, Biovia Discovery Studio (TOPKAT), CASEAR Model for Developmental Toxicity, TIMES, Leadscope Model Applier (Instem), MCASE, rtER expert system developed by EPA as well as the associated ER binding profiler as encoded within the QSAR Toolbox, Expert-based DART scheme within the OECD Toolbox, VEGA Developmental/Reproductive Toxicity library (PG, v.1.1.0). Danish (Q)SAR Database and Models free websites with predictions and models for a number of endpoints related to endocrine and molecular activity, e.g., ER, AR and TH endpoints, in the database as well as for teratogenicity from a commercial MultiCASE model, remodelled by DTU in Leadscope (now Instem) and SciQSAR by agreement from MultiCASE.

Table 32. References and context of use for prediction tools

The following table is a (non-comprehensive) list of prediction tools. This table provides easy access to resources that were mentioned or related to the topic addressed in this guidance document.

Endpoint	Name of tool	Reference	Context of use
Physicochemical properties	EPI Suite	US EPA https://www.epa.gov/tsca-screening-tools/download-epi-suitetm-estimation-program-interface-v411	Estimates parameters such as *log* K_{ow}, melting point, boiling point, vapour pressure, water solubility EPA EPI Suite™ uses a single input to run the following estimation programs: KOWWIN™, AOPWIN™, HENRYWIN™, MPBPWIN™, BIOWIN™, BioHCwin, KOCWIN™, WSKOWWIN™, WATERNT™, BCFBAF™, HYDROWIN™, KOAWIN and AEROWIN™, and the fate models WVOLWIN™, STPWIN™ and LEV3EPI™
	CompTox Dashboard	US EPA https://comptox.epa.gov/dashboard/	Multi-functional database built to meet worldwide regulatory needs (web-based platform for integration, processing, visualisation, and delivery of data/resources for a broad array of regulatory programs)
	ACD/Percepta	ACD/Labs http://www.acdlabs.com/products/percepta/	Estimator of parameters including water solubility, boiling point, *log* K_{ow}, *log D*, pK_a
	ADMET Predictor	Simulations Plus https://www.simulations-plus.com/software/admetpredictor/	Predictors for ADME, kinetics and physicochemical properties
	QSAR Toolbox	OECD https://qsartoolbox.org/	Encodes the EPISuite predictors. In addition, contains available experimental data on key physicochemical parameters
	ChemAxon	ChemAxon http://www.chemaxon.com/	Predictors for *log* K_{ow}, *log D*, pK_a

Endpoint	Name of tool	Reference	Context of use
	ChemProp	Helmholtz Centre for Environmental Research	Predictors for physicochemical properties, environmental fate, ecotoxicity, human health endpoints (ER, AR, TR, Mutagenicity, carcinogenicity, repeat dose toxicity)
	T.E.S.T.	US EPA https://www.epa.gov/chemical-research/toxicity-estimation-software-tool-test	Physical property endpoints include boiling point, flash point, surface tension, viscosity, density, water solubility, and thermal conductivity
	iSafeRat	KREATiS iSafeRat Online® https://isaferat.kreatis.eu/	Estimations parameters such as *log* K_{ow}, water solubility, vapour pressure, Henry's law constant
Aquatic toxicity	ECOSAR	US EPA https://www.epa.gov/tsca-screening-tools/ecological-structure-activity-relationships-ecosar-predictive-model	Main utility is to predict acute and chronic effects for fish, *daphnia* and algae. Structured in the form of SARs for specific chemical classes which provide some indication of likely MOA.
	Aquatic toxicity classification by ECOSAR Profiler within the QSAR Toolbox	OECD https://qsartoolbox.org/	Profiler to help assign MOA on the basis of chemical class for the purposes of deriving endpoint specific chemical categories
	CompTox Dashboard	US EPA https://comptox.epa.gov/dashboard/	Multi-functional database built to meet worldwide regulatory needs (web-based platform for integration, processing, visualisation, and delivery of data/resources for a broad array of regulatory programs)
	Biovia Discovery Studio (TOPKAT)	https://www.3ds.com/fr/produits-et-services/biovia/	Global models to predict acute toxicity to fish and daphnia. Based on Biovia Discovery Studio (TOPKAT® Toxicity Prediction by Komputer Assisted Technology) with updated training sets and advanced modeling techniques from BIOVIA Enterprise Platform®

Endpoint	Name of tool	Reference	Context of use
	T.E.S.T.	US EPA https://www.epa.gov/chemical-research/toxicity-estimation-software-tool-test	Models to predict 96-hr fathead minnow LC_{50}, 48-hr Daphnia magna LC_{50}, Tetrahymena pyriformis 50% IGC_{50}
	TIMES (TIssue MEtabolism Simulator)	LMC http://oasis-lmc.org/products/software/times.aspx	Available models include those to predict 96-hr fathead minnow LC_{50}, 48-hr *Daphnia magna* LC_{50}, *Tetrahymena pyriformis* 50% IGC_{50}
	QSAR Toolbox	OECD https://qsartoolbox.org/	Contains available experimental data to help develop new trend analysis for the prediction of key aquatic endpoints
	Verhaar rulebase; Acute aquatic classification by Verhaar profiler	JRC and OECD https://qsartoolbox.org/	The Verhaar rulebase enables substances to be characterised according to their likely MOA. The scheme has been implemented into Toxtree as well as the QSAR Toolbox
	OASIS Acute Aquatic Toxicity Mode of Action Profiler	OECD https://qsartoolbox.org/	A scheme to enable substances to be categorised according to their likely MOA.
	Lazy Structure Activity Relationships (LAZAR)		Model to predict 96 hr fathead minnow LC_{50}
	Danish (Q)SAR Database and Models	https://qsar.food.dtu.dk and https://qsarmodels.food.dtu.dk/	Predictions for 650,000 substances and models to predict short-term toxicity to Fathead minnow (96h LC_{50}), *Daphnia magna* (48h LC_{50}), and *Pseudokirchneriella subcapitata* (72h EC_{50}) in the free websites
	iSafeRat	KREATiS iSafeRat Online® https://isaferat.kreatis.eu/	Available models include those to prediction 96-hr fish LC_{50}, 48-hr daphnids EC_{50}, 72-hr algae ErC_{50}, 32-d fish EC_{10}, 21d daphnids EC_{10}, 72-hr algae ErC_{10}, ASRIT micro-organisms 30-180min EC_{50}
Biodegradation	BIOWIN	US EPA	The BIOWIN program in US EPA's EPI Suite™ estimates the

Endpoint	Name of tool	Reference	Context of use
		https://www.epa.gov/tsca-screening-tools/download-epi-suitetm-estimation-program-interface-v411	probability of rapid aerobic and anaerobic biodegradation of an organic compound in the presence of mixed populations of environmental microorganisms. It contains seven models: Biowin1 (linear probability), Biowin2 (nonlinear probability), Biowin3 (expert survey ultimate biodegradation), Biowin4 (expert survey primary biodegradation), Biowin5 (MITI linear), Biowin6 (MITI nonlinear), Biowin7 (anaerobic biodegradation)
	CompTox Dashboard	US EPA https://comptox.epa.gov/dashboard/	Multi-functional database built to meet worldwide regulatory needs (web-based platform for integration, processing, visualisation, and delivery of data/resources for a broad array of regulatory programs)
	Biovia Discovery Studio (TOPKAT)	https://www.3ds.com/fr/produits-et-services/biovia/	Global model to predict ready biodegradability
	Catalogic	LMC http://oasis-lmc.org/products/software/catalogic.aspx	Contains a suite of models to predict biodegradation under different study protocols e.g. OECD 301C While accounting for microbial metabolism
	Danish (Q)SAR Database and Models	https://qsar.food.dtu.dk and https://qsarmodels.food.dtu.dk/	Predictions for 650,000 substances and models to predict Not ready biodegradability in the free websites
	ChemProp	Helmholtz Centre for Environmental Research	Predictors for physicochemical properties, environmental fate, ecotoxicity, human health endpoints (ER, AR, TR, Mutagenicity, carcinogenicity, repeat dose toxicity)
Bioaccumulation	Catalogic	LMC	Contains BCF base line model which predicts BCF While accounting for modulating

Endpoint	Name of tool	Reference	Context of use
		http://oasis-lmc.org/products/software/catalogic.aspx	factors such as metabolism, size, ionisation
	CompTox Dashboard	US EPA https://comptox.epa.gov/dashboard/	Multi-functional database built to meet worldwide regulatory needs (web-based platform for integration, processing, visualisation, and delivery of data/resources for a broad array of regulatory programs)
	BCFBAF	US EPA https://www.epa.gov/tsca-screening-tools/download-epi-suitetm-estimation-program-interface-v411	The BCFBAF program in EPA's EPI Suite™ estimates BCF of an organic compound using the compound's *log* K_{ow}. BCFBAF includes estimation of the Biotransformation Rate (kM) in fish and estimation of Bioaccumulation Factor (BAF) by the Arnot-Gobas method (2003)
	CAESAR Model for BCF	CAESAR https://www.vegahub.eu/	Global model for BCF, now part of the VEGA platform
	T.E.S.T.	US EPA https://www.epa.gov/chemical-research/toxicity-estimation-software-tool-test	Software encodes the CAESAR BCF model
	iSafeRat	KREATiS iSafeRat Online® https://isaferat.kreatis.eu/ https://isaferat.kreatis.eu/	Contains a base line model which predicts the BCF
Mammalian toxicity: Acute toxicity	CompTox Dashboard	US EPA https://comptox.epa.gov/dashboard/	Multi-functional database built to meet worldwide regulatory needs (web-based platform for integration, processing, visualisation, and delivery of data/resources for a broad array of regulatory programs)
	T.E.S.T.	US EPA https://www.epa.gov/chemical-research/toxicity-estimation-software-tool-test	Global model for the prediction of rat LD_{50}

Endpoint	Name of tool	Reference	Context of use
	Biovia Discovery Studio (TOPKAT)	https://www.3ds.com/fr/produits-et-services/biovia/	Global model for the prediction of rat LD_{50}
	QSAR Toolbox	OECD https://qsartoolbox.org/	Contains experimental data of LD_{50} in rodents
	Danish (Q)SAR Database	https://qsar.food.dtu.dk	Predictions for 650,000 substances from ACD/Labs models for Rat and Mouse LD_{50} by different administration routes in the free website
	TIMES (TIssue MEtabolism Simulator)	LMC https://oasis-lmc.org/products/models/human-health-endpoints/acute-oral-toxicity.aspx	The acute oral toxicity model predicts the median lethal dose of a substance that causes toxic effect to 50% (LD_{50}) of the test rodent (rat or mouse) within a designated period
Mammalian toxicity-eye	Biovia Discovery Studio (TOPKAT)	https://www.3ds.com/fr/produits-et-services/biovia/	Global model to discriminate eye irritation potency
	BFR Rulebase within Toxtree	JRC	Scheme to classify eye irritants on the basis of structural alerts and to assign no classification for substances that meet specific physicochemical parameter thresholds
	Eye irritation inclusion and exclusion rules by BfR Profiler within the QSAR Toolbox	OECD https://qsartoolbox.org/	Permits a categorisation of substances based on presence of alerts and extremes of physicochemical parameters
	Derek Nexus	Lhasa Limited https://www.lhasalimited.org/solutions/	Contains SARs for eye irritation that is useful to characterise MOA information
	iSafeRat	KREATiS iSafeRat Online® https://isaferat.kreatis.eu/	Includes a categorical model to predict if the chemical falls within the corrosive, irritant or non-irritant category
	TIMES (TIssue MEtabolism Simulator)	LMC	Includes an eye irritation model

Endpoint	Name of tool	Reference	Context of use
		https://oasis-lmc.org/products/software/times.aspx	
Mammalian toxicity: Skin irritation	Danish (Q)SAR Database	https://qsar.food.dtu.dk and https://qsarmodels.food.dtu.dk/	Predictions for 650,000 substances and models for severe versus mild/no skin irritation in the free websites
	Derek Nexus	Lhasa Limited https://www.lhasalimited.org/solutions/	Contains SARs for skin irritation that is useful to characterise MOA information
	iSafeRat	KREATiS iSafeRat Online® https://isaferat.kreatis.eu/	Includes a categorical model to predict if the chemical falls within the corrosive, irritant or non-irritant category
	TIMES (TIssue MEtabolism Simulator)	LMC https://oasis-lmc.org/products/models/human-health-endpoints/skin-irritationcorrosion.aspx	The model predicts the reversible (irritation) and irreversible (corrosion) damage of the skin following the application of a test substance for up to 4 hours
Mammalian toxicity: skin sensitisation	CAESAR Model for Skin Sensitisation	CAESAR https://www.vegahub.eu/	Global model for sensitisation, now part of the VEGA platform
	TIMES (TIssue MEtabolism Simulator)	LMC https://oasis-lmc.org/products/software/times.aspx	Hybrid expert system to predict sensitisation potency While accounting for metabolism
	Biovia Discovery Studio (TOPKAT)	https://www.3ds.com/fr/produits-et-services/biovia/	Global model for the prediction of sensitising potency
	Derek Nexus	Lhasa Limited https://www.lhasalimited.org/solutions/	Contains SARs for sensitisation which are helpful to assign MOA Contains high potency category SARs for use in assessing Dermal Sensitisation Thresholds
	CASE Ultra	Multicase Inc. http://multicase.com/case-ultra	Global model for prediction of sensitisation

Endpoint	Name of tool	Reference	Context of use
	Protein binding by OECD profiler within the QSAR Toolbox	OECD https://qsartoolbox.org/	Alerts which characterise electrophilic reactivity based on organic chemistry principles. Assigns substances into reaction mechanistic domains that are pertinent for the assessment of skin sensitisation potential
	Protein binding by OASIS profiler within the QSAR Toolbox Protein binding alerts for skin sensitisation by OASIS within the QSAR Toolbox	OECD https://qsartoolbox.org/	Mirror the SARs contained within the TIMES skin sensitisation model
	SMARTS alerts within Toxtree	JRC	Assignment of reaction mechanistic domains. Based on the reaction principles defined by Aptula and Roberts, (2006)
	Danish (Q)SAR Database	https://qsar.food.dtu.dk	Predictions for 650,000 substances from MultiCASE model for allergic contact dermatitis, and from models by DTU on the same training set in agreement with MultiCASE
	iSafeRat	KREATiS iSafeRat Online® https://isaferat.kreatis.eu/	Includes a model which predicts if the chemical falls within the sensitizer or not sensitizer category and predicts the potency of the chemical based on the estimation of a concentration for sensitisation
Mammalian toxicity: mutagenicity/genotoxicity	CAESAR	CAESAR https://www.vegahub.eu/	Global model for Ames mutagenicity, now part of the VEGA platform
	TIMES (TIssue MEtabolism Simulator)	LMC https://oasis-lmc.org/products/software/times.aspx	Hybrid expert system which contains a suite of models for the prediction of Ames mutagenicity, *in vitro* chromosomal aberration, *in vivo* liver genotoxicity and *in vivo* micronucleus. All models account for metabolism

Endpoint	Name of tool	Reference	Context of use
	Biovia Discovery Studio (TOPKAT)	https://www.3ds.com/fr/produits-et-services/biovia/	Global model for Ames mutagenicity
	Derek Nexus	Lhasa Limited https://www.lhasalimited.org/solutions/	SAR for mutagenicity, chromosomal aberration, DNA damage etc Useful to categorise chemicals on the basis of their likely MOA
	Sarah Nexus	Lhasa Limited https://www.lhasalimited.org/solutions/in-silico-mutagenicity-assessment/	Statistical system, Global model for Ames mutagenicity
	Acrostic	Lhasa Limited https://www.lhasalimited.org/solutions/in-silico-mutagenicity-assessment/	Tool applying read-across methodology for nitrosamine mutagenicity
	META Ultra	Multicase Inc. http://multicase.com/meta-ultra	Global model for Ames mutagenicity
	Leadscope Model Applier (Instem)	Instem (formerly Leadscope) https://www.instem.com/solutions/	Global models for a range of different genetic toxicity endpoints
	DNA binding by OASIS profiler within the QSAR Toolbox DNA Alerts for Ames, CA and MN by OASIS within the QSAR Toolbox	OECD https://qsartoolbox.org/	Mirror the SARs contained within the TIMES
	DNA Binding by OECD profiler within the QSAR Toolbox	https://qsartoolbox.org/	Alerts which characterise electrophilic reactivity based on organic chemistry principles
	Benigni/Bossa Rulebase within Toxtree	JRC	A compilation of SARs made by R Benigni and C Bossa
	ToxMIC-ISS plug-in allows the identification of Structure Alerts for the *in vivo* micronucleus assay within Toxtree	JRC	A compilation of SARs for *in vivo* MN

Endpoint	Name of tool	Reference	Context of use
	In vitro, *in vivo* and Carcinogenicity/mutagenicity alerts by ISS within the QSAR Toolbox	OECD https://qsartoolbox.org/	An update and refinement of the Benigni-Bossa and Toxx-MIC rulebases re-coded within the Toolbox
	ChemProp	Helmholtz Centre for Environmental Research	Predictors for Physicochemical properties, environmental fate, ecotoxicity, human health endpoints (ER, AR, TR, Mutagenicity, carcinogenicity, repeat dose toxicity)
	T.E.S.T.	US EPA https://www.epa.gov/chemical-research/toxicity-estimation-software-tool-test	Global model for Ames mutagenicity
	Lazy Structure Activity Relationships (LAZAR)		Global models based on the following datasets DSSTox Carcinogenic Potency DBS and Kazius-Bursi Salmonella
	Danish (Q)SAR Database	https://qsar.food.dtu.dk and https://qsarmodels.food.dtu.dk	Predictions for 650,000 substances and models (non-commercial/non-confidential models only) for Ashby fragments (commercial, only available in the database) *In vitro* mutagenicity: • Ames test • Ames sub-models (DTU models for S9 activation, Base pair mutation, Frame shift mutation, Potency at least 10 times over control group based on P&G confidential data, only available in the database) • Chromosomal aberration CHO (commercial model, only available in the database) • Chromosomal aberration CHL • Mouse lymphoma TK assay • Unscheduled DNA synthesis • CHO/HGPRT forward mutation assay • SHE cell transformation

Endpoint	Name of tool	Reference	Context of use
			In vivo mutagenicity: • *Drosophila* sex-linked recessive lethal • Mouse micronucleus bone marrow • Rodent dominant lethal • Mouse SCE bone marrow • Mouse Comet assay
Mammalian toxicity: repeated dose toxicity	Biovia Discovery Studio (TOPKAT) (LOAEL, MTD)	https://www.3ds.com/fr/produits-et-services/biovia/	Global model for the prediction of LOAEL, MTD
	Derek Nexus	Lhasa Limited https://www.lhasalimited.org/solutions/	SARs for many different endpoints associated with repeated dose toxicity
	ChemProp	Helmholtz Centre for Environmental Research	Predictors for Physicochemical properties, environmental fate, ecotoxicity, human health endpoints (ER, AR, TR, Mutagenicity, carcinogenicity, repeat dose toxicity)
	QSAR Toolbox	OECD https://qsartoolbox.org/	Predictors for Physicochemical properties, environmental fate, ecotoxicity, human health endpoints (ER, AR, TR, Mutagenicity, carcinogenicity, repeat dose toxicity)
	Fraunhofer database	http://www.fraunhofer-repdose.de/	Database of repeated dose toxicity information. Also made available within the QSAR Toolbox
	Lazy Structure Activity Relationships (LAZAR)		Global FDA v3b Maximum Recommended Daily Dose model
	Hazard Evaluation Support System Integrated Platform (HESS) Profiler	OECD https://qsartoolbox.org/	Profiler within the QSAR Toolbox to help assign MOA
	Hazard Evaluation Support System Integrated Platform (HESS)	National Institute of Technology and Evaluation (NITE) https://www.nite.go.jp/en/chem/qsar/hess-e.html	Expert system containing repeated dose toxicity information to facilitate hazard assessment through the

Endpoint	Name of tool	Reference	Context of use
			development of chemical categories. System mimics the structure/platform of the QSAR Toolbox
Mammalian toxicity: reproductive and developmental toxicity	Derek Nexus	Lhasa Limited https://www.lhasalimited.org/solutions/	SARs for teratogenicity, developmental toxicity, reproductive effects
	Biovia Discovery Studio (TOPKAT)	https://www.3ds.com/fr/produits-et-services/biovia/	Global model for developmental toxicity
	CAESAR Model for Developmental Toxicity	CAESAR https://www.vegahub.eu/	Global model for developmental toxicity
	Leadscope Model Applier (Instem)	Instem (formerly Leadscope) https://www.instem.com/solutions/	Sex specific global models for developmental toxicity and reproductive effects in rodents
	META Ultra	Multicase Inc. http://multicase.com/meta-ultra	
	Expert-based Developmental and Reproductive Toxicity (DART) scheme within the QSAR Toolbox	OECD https://qsartoolbox.org	Profiler within the QSAR Toolbox to help assign MOA
	rtER expert system developed by EPA as well as the associated ER binding profiler within the QSAR Toolbox	OECD https://qsartoolbox.org/	Encoded as an expert system within the QSAR Toolbox and as a profiler to assign chemicals on the basis of their likely ER MOA
	Oasis TIMES	http://oasis-lmc.org/products/software/times.aspx	Models for ER, and AR Binding affinities
	Danish (Q)SAR Database and Models	https://qsar.food.dtu.dk and https://qsarmodels.food.dtu.dk	Predictions for 650,000 substances from a number of endpoints related to endocrine and molecular activity, e.g. ER, AR and TH endpoints, as well as for teratogenicity from a commercial MultiCASE model, remodelled by DTU in Leadscope and SciQSAR by agreement from MultiCASE

Endpoint	Name of tool	Reference	Context of use
Potential for metabolisation and potential metabolites	QSAR TB (skin, liver, environmental/ simulated or observed, / metabolites indicated but not probability/ freely downloadable from OECD web site) SMARTcyp: predicts the sites in molecules that are most liable to cytochrome P450 mediated metabolism. Probability for reaction with CYPs Meta2print: As above & Identity and probability of metabolites generated, based on both phase I and II reactions Oasis TIMES BioTransformer	https://sourceforge.net/projects/metaprint2d/ University of Cambridge/ Department of Chemistry & Unilever Centre for Molecular Science Informatics	Various in relation to when potential metabolisation is of significance for chemical categorisation.
	CompTox Dashboard	US EPA https://comptox.epa.gov/dashboard/	Multi-functional database built to meet worldwide regulatory needs (web-based platform for integration, processing, visualisation, and delivery of data/resources for a broad array of regulatory programs).
	Danish (Q)SAR Database	https://qsar.food.dtu.dk and https://qsarmodels.food.dtu.dk	Predictions for 650,000 substances and models for CYP2D6, CYP2C9, PXR, AhR and CAR endpoints.

Grouping and read-across reporting template

An example reporting template is attached to this document (as a Word document).

The template is intended to guide the reporting of analogue and category approaches, as described in Section 7 of this document, reflecting the essential elements for documenting the results and reasoning of the grouping and read-across.

Data matrix template

Example templates for building a data matrix for the analogue and category approach are attached to this document (as an Excel spreadsheet).

and in:

- Table 26. Example data matrix, analogue approach
- Table 29. Example data matrix, chemical category

Reporting of omics data for chemical grouping

General guidance on reporting of omics data from laboratory-based toxicology studies can be found in the OECD Omics Reporting Framework (OORF) (OECD 2023a).

Specific guidance related to reporting omics data in the context of grouping is provided in the Chemical Grouping Application Reporting Module (CG-ARM), https://one.oecd.org/document/ENV/CBC/MONO(2025)20/en/pdf.

Glossary of selected terms

Adverse Outcome Pathway (AOP): an AOP describes a sequence of events commencing with initial interaction(s) of a stressor with a biomolecule within an organism that causes a perturbation in its biology.

Allowed differences: the degree of (e.g. structural, physicochemical, biological) differences considered acceptable among members of a group or category, not affecting the endpoint of interest. This may depend on the problem formulation, available data, and regulatory decision context.

Analogue (or analogous substance): is one substance that has been identified as exhibiting similarity (see definition) to another substance.

Analogue approach: an approach where the available data for typically one substance ("the source") can be used to predict the outcome of a second similar (see definition) substance ("the target").

Applicability domain: the Applicability Domain (AD) of a (Q)SAR model, as described in the Guidance Document (OECD, 2007), is the response and chemical structure space in which the model makes predictions with a given reliability. Elaborating on the AD definition given above, there is no consensus on how to define an AD, however the parametric, structural, mechanistic, metabolic and response space of the model can be considered in the definition.

In the context of read-across, the set of required structural elements and allowed structural differences at minimum define the applicability domain (or boundaries) of the category and define which substances can be part of the category and which are not (category membership).

Bioactivity similarity: describes the similarity(ies) between two or more omics profiles (or signatures), where those profiles (or signatures) are measured in a defined test system following exposure to two or more test substances, often as part of a bridging study.

Boundary chemical: category members falling at the opposite extremes of a trend and between which interpolations are considered reliable are called boundary chemicals.

Breakpoint chemical: a point of discontinuity, change, or cessation. A chemical that identifies a turning point in a trend is called a breakpoint chemical.

Bridging studies: bridging studies are defined as but not limited to studies conducted to show relevance or create a bridge between the substances included in an analogue or category approach to establish the similarity in properties, (eco)toxicological profile, and/or environmental fate and behaviour.

Category: group of substances (or category members) that are defined based on similarity. Category substances (or members) require supporting data to justify their inclusion in the category, termed the category justification.

***Category approach*:** an approach where the available data for two or more substances can be used to predict the outcome of one or more similar (see definition) substances.

Conventional toxicology: traditional methods and approaches used to assess the potential toxicity of substances, particularly in contrast to NAMs.

Defined Approaches (DAs): a formalised decision-making approach consisting of a fixed data interpretation procedure used to interpret data from a defined set of information elements, (OECD, 2020a).

Endpoint: any single or group of physicochemical, biological, or environmental properties that can be measured/modelled. An endpoint could be determined by different experimental protocols and under different experimental conditions.

Expert system:[140] a formalised system, usually computerised that enables an end-user to make rational predictions of toxicity based on structure alone. Expert systems are typically categorised by whether they are underpinned by empirically based algorithms such as QSARs e.g. T.E.S.T., OPERA; knowledge bases such as SARs e.g. Derek Nexus, Toxtree or a hybrid of the two e.g. TIMES, ChemTunes.

Extrapolation: the estimation of a value for a member that is near or at the category boundary using measured values from internal category members.

Features: genes, proteins and/or endogenous metabolites that can be annotated or unannotated, and which together form input data for a bioactivity similarity assessment.

***Group representative chemicals*:** category members falling towards the centre of a trend, and which could be used to represent the group, e.g. as a prioritisation strategy for higher tier testing, are called representative chemicals.

'Grouping' or 'Chemical grouping': the process of identifying a collection of substances that are likely to be similar or follow a regular pattern as a result of similarity (see definition).

Integrated Approaches to Testing and Assessment (IATA): see Appendix 1 in OECD GD 329 IATA-'Overview of Concepts and Available Guidance related to Integrated Approaches to Testing and Assessment (IATA)', (OECD, 2020a).

Interpolation: 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.

Metabolism: a linked series of chemical reactions in the body to convert a chemical (i.e. a xenobiotic) to either an inactive compound or to a more active compound for excretion from the body. These chemical reactions form metabolites, which can also be referred to as biotransformation products.

Mode of Action: a biologically plausible sequence of key events at different levels of biological organisation, starting with the exposure to a chemical and leading to an observed (adverse) effect, (OECD, 2020a).

Mechanism of Action: a detailed molecular description of the mechanistic interaction through which a substance/molecule produces its effect, (OECD, 2020a).

Moiety:[141] in physical organic chemistry moiety is generally used to signify part of a molecule, e.g. in an ester R1COOR2 the alcohol moiety is R2O. The term should not be used for a small fragment of a molecule.

Omics: in the context of this guidance document, omics refers to technologies that are used to measure a broad range of molecular responses to chemical exposure. Widely used approaches include transcriptomics (study of expression of multiple genes) and metabolomics (study of levels of multiple endogenous metabolites and the biochemical processes in which they are involved in), within a cell, tissue or organism.

[140] Dearden, J.C., Barratt, M.D., Benigni, R., Bristol, D.W., Combes, R.D., Cronin, M.T.D., Judson, P.N., Payne, M.P., Richard, A.M., Tichý, M., Worth, A.P. and Yourick, J.J., The development and validation of expert systems for predicting toxicity. ATLA, 25 (1997) 223–252.

[141] IUPAC Compendium of Chemical Terminology, 3rd ed. International Union of Pure and Applied Chemistry; 2006. Online version 3.0.1, 2019. https://doi.org/10.1351/goldbook.M03968

Omics profile: comprises the set of all measured features or a subset of statistically pre-filtered features, i.e. this is a data-driven profile that does not use any external toxicological knowledge.

Omics signature: comprises a pre-specified, reduced (i.e. targeted) set of measured features that are associated with one (or more) molecular pathway, MOA, AOP or endpoint, i.e. this is a knowledge driven signature that does use external toxicological knowledge.

***Quantitative Structure Activity Relationship, QSAR model*:** a QSAR is a mathematical model (often a statistical correlation) relating one or more quantitative parameters derived from chemical structure to a quantitative measure of a property or activity (e.g. a (eco)toxicological endpoint). QSARs are quantitative models yielding a continuous or categorical result.

(Quantitative) Structure Activity Relationship, (Q)SAR model: SARs and QSARs, collectively referred to as (Q)SARs, are theoretical models that predict the (quantitative) value of a property of a substance using as input information on the structure. In this document, expert and rule-based systems are considered part of this definition.

Read-across: a technique for predicting endpoint information for one substance (target substance), by using data from the same endpoint from (a) similar substance(s), (source substance(s)).

Similarity: several factors should be considered when evaluating similarity. These factors can include structure, physicochemical properties, chemical reactivity profile, bioactivity, conventional toxicological profile and ADME/TK including metabolism.

Source substance (or source analogue): a chemical that has been identified as an appropriate chemical for use in a read-across based on similarity to the target chemical and existence of relevant data.

Structure Activity Relationship (SAR): qualitative relationship (i.e. an association) between a molecular (sub)structure and the presence or absence of a biological activity, or the capacity to modulate a biological activity imparted by another substructure. A substructure associated with the presence of a biological activity is sometimes called a structural alert.

Target substance: substance of interest for which data gaps exist that need to be addressed**.**

Toxic Equivalents (TEQ) approach: the most toxicologically relevant compound is used as the reference compound. This compound does not necessarily have to be present in the mixture being assessed, but the components of the mixture must all act by the same single toxic pathway and be of the same compound type (structural/functional group similarity) as the reference.

Uncertainty: a general term referring to all types of limitations in available knowledge that affect the range and probability of possible answers to an assessment question (EFSA, 2018a).

Weight of Evidence (WoE): a stepwise process/approach of collecting and weighing evidence to reach a conclusion on a particular problem formulation including assessment of the degree of confidence (OECD, 2019a).

Guidance on Grouping of Chemicals, Third Edition

Series on Testing and Assessment

The OECD Guidance on Grouping of Chemicals describes approaches to consider closely related chemicals as a group rather than individually and apply read-across approaches for filling data gaps. The first edition was published in 2007 and the second edition was published in 2014. This third edition incorporates New Approach Methodologies (NAMs) such as adverse outcome pathways (AOPs), omics technologies, high-throughput screening and (Q)SAR models for developing groups and substantiating similarity. The guidance outlines how to establish the scientific basis for read-across approaches to fill data gaps, integrates insights gained from the OECD IATA Case Study Project and describes new methods to quantify read-across performance and uncertainties. This edition expands coverage of mechanistic-based approaches and includes guidance for metals and nanomaterials grouping. The document provides critical guidance for assessing the hazards of chemicals using approaches that gain efficiencies and reduce reliance on animal testing.