ECHA
EUROPEAN CHEMICALS AGENCY

Non-animal approaches

Current status of regulatory applicability under the REACH, CLP and Biocidal Products regulations

Report
November 2017

Legal Notice

Non-animal approaches - Current status of regulatory applicability under the REACH, CLP and Biocidal Products regulations

Reference: ECHA-17/R/24/EN
Cat. number: ED-05-17-096-EN-N
ISBN: 978-92-9020-208-0
DOI: 10.2823/000784
Date: November 2017
Language: English

If you have questions or comments in relation to this document please send them (quote the reference and issue date) using the information request form. The form can be accessed from the "Contact ECHA" page at: http://echa.europa.eu/contact

European Chemicals Agency
Mailing address: P.O. Box 400, FI-00121 Helsinki, Finland
Visiting address: Annankatu 18, Helsinki, Finland

Version	Changes
1.0 (November 2017)	First edition

Table of Contents

Foreword

Dear reader,

This is the first time we report on regulatory applicability of non-animal approaches, describing how currently available reliable non-animal approaches can be used under the REACH Regulation and for active substances under the Biocidal Products Regulation (BPR) to fulfil the information requirements and reflecting the needs of the Classification, Labelling and Packaging (CLP) Regulation. The report was requested by ECHA's Management Board.

Non-animal approaches for investigating hazardous properties of substances are increasingly used and promoted. Vertebrate animal testing should be the last resort, only used after all other scientifically reliable methods have been exhaustively explored.

We have reported regularly to the European Commission on how companies use non-animal approaches under the REACH Regulation (*The use of alternative to testing on animals for the REACH Regulation*: 2011, 2014 and 2017). These reports have confirmed that the most important tool to avoid unnecessary animal testing, i.e. data sharing, is working well. The duty holders also use extensively adaptation methods and non-animal approaches such as read-across, weight of evidence, computer modelling and *in vitro* methods. However, in many cases the quality of information used by registrants in adaptations and from non-animal approaches submitted is not robust enough to comply with the legal criteria and thus to replace vertebrate testing. We encourage duty holders to use reliable non-animal approaches and have provided tools and advice to industry to support compliance with the legal requirements. We have updated our guidance, practical guides, published case studies and webinars.

To further improve understanding on how the non-animal approaches can be used to meet the legal requirements, we describe in this report the main types of relevant non-animal approaches and adaptation rules (Part A – General considerations) and the information requirements and needs for relevant (eco)toxicological properties (Part B – Specific considerations). In Part B, we describe for each relevant information requirement the potential approaches utilising non-animal approaches, the challenges to achieving this and future perspectives, including those approaches that could be close to regulatory applicability. Duty holders can use this report in conjunction with ECHA Guidance on non-animal approaches in their consideration and preparation of dossiers.

During the preparation of the report, various relevant external agencies and institutions, as well as ECHA's accredited stakeholders, have been consulted and their feedback collected to ensure that all the relevant non-animal approaches are considered and given the right perspective (see the list of contributing partners in Appendix 1).

I hope that this report indicates to the scientific community where further scientific development and/or regulatory acceptance is needed to reduce animal testing in implementing the three regulations that ECHA coordinates. I also hope this report stimulates consensus among authorities and stakeholders about the opportunities and current limitations related to non-animal approaches. We have included general recommendations in the executive summary which reflect the complex interplay between the different legal information requirements of REACH, CLP and the BPR.

I hope you find the report to be of interest to you.

Geert Dancet,

Executive Director

Executive summary

Significant developments have taken place over the last decade to replace vertebrate animal testing for chemical safety assessment with non-animal approaches. Scientific research has led to further progress with concepts, such as the integrated approach to testing and assessment (IATA) and adverse outcome pathway (AOP) frameworks, which can be used to integrate data from various non-animal approaches into biologically justifiable steps of investigations to inform on (eco)toxicological properties. These developments have already resulted in a reduced need for animal testing under the EU chemicals legislation and hold promise for enabling the regulators to take further steps towards a stronger implementation of the 3Rs principle of replacement, reduction, and refinement of animal testing.

In light of the above advances, a careful review and analysis of how the latest scientific knowledge and developments could be transferred into regulatory use are necessary. This report reviews the current status of the regulatory applicability of non-animal approaches and how they may inform on the need for hazard classification and be used to fulfil the regulatory information requirements for the purpose of chemical safety assessments. Our focus here is on three EU regulations: the REACH Regulation, the Classification, Labelling and Packaging (CLP) Regulation, and the Biocidal Products Regulation (BPR). The provisions and requirements of these regulations still largely rely on the use of vertebrate animal testing. At the same time, they all involve the obligation to use vertebrate animal testing as a last resort and, therefore, provide significant opportunities for using non-animal approaches, where appropriate.

The report focuses not only on non-animal approaches as a direct replacement of animal testing, but also on how they can be used together with other data as supporting evidence in regulatory processes, and what reduction and refinement methods are available when animal testing cannot be avoided. The report also provides an outlook of the near and medium-term future expectations and identifies ways and means to accelerate the development of non-animal approaches, enhance their applicability and promote their wider use.

The report is aimed at a wide audience including Member State competent authorities, the research community and other stakeholders, such as registrants under the REACH Regulation and applicants under the BPR.

As a result of significant international collaboration and progress made in the development of non-animal approaches, the regulatory acceptance of many non-animal approaches has been achieved for some of the so-called lower-tier information requirements. For example, the information requirements on skin corrosion/irritation, serious eye damage/eye irritation, and skin sensitisation can now be fulfilled by applying non-animal approaches such as *in vitro* tests, IATA and AOP concepts. Non-animal approaches for these toxicological properties are already defined in the legislation as the default method to generate the information in most cases.

While the standard information requirements rely on animal testing for most endpoints, the legislation highlights that animal testing is a last resort. It also allows the fulfilment of the standard information requirements through the so-called adaptations, which include the use of grouping and read-across, and the use of weight-of-evidence. Information from non-animal approaches may be used as supporting data for a grouping and read-across adaptation or as elements of a weight-of-evidence adaptation. For instance, a specific weight-of-evidence[1] approach may be used for substances with a low sub-acute oral toxicity to adapt the requirement for an acute oral toxicity study and avoid animal testing for this endpoint in some cases. In general, adaptation of the standard *in vivo* studies works better for lower-tier

[1] Defined as information combined and weighed from several pieces of evidence.

endpoints, as toxic mechanisms are less complex and therefore easier to predict with non-animal approaches (e.g. QSAR, *in vitro* tests) than for higher-tier endpoints. However, further promotion of the available non-animal approaches for lower-tier studies and awareness-raising are still needed to ensure their application.

The legislation stipulates that the data generated must be adequate for hazard classification and labelling and risk assessment. This is key to ensure that the protection goals for human health and the environment as laid down in the Regulations are not compromised.

In practice, the results from (eco)toxicity testing are used for hazard classification in line with the criteria set in the CLP Regulation. This leads to labelling of hazardous chemicals with warning symbols and instructions for safe use. The type of information needed for hazard classification depends on the hazard class and its (sub)categories and may be based on:

- the dose or concentration levels causing severe effects (for acute toxicity, skin sensitisation and specific target organ toxicity);
- the severity of toxic effects (for skin corrosion/irritation and eye damage/eye irritation); or
- the strength of the evidence demonstrating toxic effects (for germ cell mutagenicity, carcinogenicity and reproductive toxicity).

Generally, the classification criteria refer mainly to human and animal data, which poses a challenge for the comparison of non-animal data with these criteria.

The test results are also used in risk assessments under the REACH Regulation and the BPR, as they provide the points of departure for determining safe levels of exposure to substances. Under the REACH Regulation, these safe level values are called derived no-effect levels (DNELs), to protect humans, and predicted no-effect concentrations (PNECs), to protect the environment. Under the BPR, acceptable exposure levels (AELs) are determined. While the derivation of safe exposure levels can be achieved with non-animal data for some local effects, this has proved to be more problematic for systemic effects. This can be due to the lack of quantitative information generated by some non-animal approaches, or the difficulty in extrapolating results from non-animal approaches to *in vivo* values, especially for higher-tier endpoints, which is necessary for deriving threshold values.

Apart from fulfilling the REACH and BPR information requirements and supporting hazard classification under CLP, non-animal approaches can provide useful information on (eco)toxicity mechanisms, bioavailability, and internal and external exposure, which can be used for screening and (de)prioritisation of substances for further regulatory action.

For higher-tier endpoints, specific non-animal approaches that could directly replace vertebrate animal tests are not yet available and not foreseen in the near or even medium-term future, and adaptations are currently the main approaches to reduce the need for new animal testing.

In spite of very active ongoing research in the area of non-animal approaches, approaches capable of replacing animal testing for complex endpoints are not yet available. Also the nature of such future approaches cannot be established yet. Furthermore, they may not provide the same level of information on the toxicity of substances as the current animal studies, for instance in terms of dose/concentration-response relationship and adverse effects. New non-animal approaches, such as *in vitro* microsystems and high-throughput/high-content methods, are under development and aim at providing more comprehensive insights into the mechanisms of toxicity than current non-animal approaches. However, they will require further standardisation, especially for the interpretation of the results and validation, before they can be considered for regulatory purposes. For example, it would need to be clarified how to make use of the evidence from new non-animal approaches that do not directly inform on adversity or specific toxicities for classification under the CLP Regulation, or how to use such data for

determining (no-)effect levels in risk assessment. Therefore, the wider use of non-animal approaches for regulatory purposes would require an extensive discussion on how results from non-animal approaches can be used in a regulatory context where classification and risk assessments are currently strongly dependent on information on effects in humans or animals. A continuous and active dialogue between the research community and regulatory authorities is also needed to ensure the regulatory relevance of the research efforts and to avoid delays in the transfer of scientific developments into regulatory use.

Building an inventory of non-animal approaches and available models predicting different types of effects, and being at different stages of development and regulatory applicability, would clarify their diversity and interrelations, and hence could facilitate their further development and application. Applying information from humans, animal studies and non-animal approaches (for example, "-*omics*") together or in parallel would enhance the interpretation of the results of non-animal approaches and the understanding of their predictive capacity and their performance, thereby increasing confidence in their ability to produce reliable and consistent results that are relevant for regulatory decision making.

ECHA stays committed to promoting the development and use of non-animal approaches. In this regard, ECHA:

- closely follows the scientific developments in the field and collaborates internationally to explore the best practice, using the information from non-animal approaches whenever applicable;
- contributes to the discussion in international fora on how the information from non-animal approaches could be applied in the context of hazard classification;
- regularly updates both its guidance and its website to include information on updated and new test methods, including non-animal approaches, thereby supporting duty holders in understanding how these methods may be used to meet the information requirements under the different regulations;
- raises awareness and promotes the newly-developed methods and approaches among its own staff, scientific committees and stakeholders; and
- gives proactive scientific and technical advice to the European Commission on specific non-animal approaches, including recommendations on the update of the corresponding legal text, and on generic aspects related to non-animal approaches.

List of abbreviations

ADME	Absorption, distribution, metabolism and excretion
AEL	Acceptable operator exposure level
AFT	Acute fish toxicity
AOP	Adverse outcome pathway
ATE	Acute toxicity estimates
AUC	Area under the curve
BCF	Bioconcentration factor
BCOP	Bovine corneal opacity and permeability assay
BMF	Biomagnification factor
BPR	Biocidal products regulation
BSAF	Biota-sediment accumulation factor
CLP	Classification, labelling and packaging
Cmax	Maximum concentration in the blood or plasma
CRED	Criteria for reporting and evaluating ecotoxicity data
CSA	Chemical safety assessment
CSR	Chemical safety report
CTA	Cell transformation assay
CYP	Cytochrome P450
DIPs	Data interpretation procedure
DNA	Deoxyribonucleic acid
DNEL	Derived no-effect level
EC_{50}	Half-maximal effective concentration (concentration inducing response halfway between the baseline and maximum after a specified exposure time)
EOGRTS	Extended one-generation reproductive toxicity study
EU-NETVAL	European Union network of laboratories for the validation of alternative methods
ESAC	EURL ECVAM scientific advisory committee
EURL ECVAM	European Union reference laboratory for alternatives to animal testing
FET	Fish embryo toxicity
GD	Guidance document
GIVIMP	Good *in vitro* method practice
GLP	Good laboratory practice
hESC	Human embryonic stem cell
hrER	Human recombinant estrogen receptor
HTS/HTPS	High-throughput screening
HTTK	High-throughput toxicokinetics
IATA	Integrated approach to testing and assessment
ICATM	International cooperation on alternative test methods
ITS	Integrated testing strategy
IUCLID	International uniform chemical information database
IVIVE	*In vitro* to *in vivo* extrapolation
KE	Key event
K_{ow}	Octanol-water partition coefficient

LC_{50}	Lethal concentration 50 (acute toxic concentration causing the death of 50 % of the test population)
LD_{50}	Lethal dose 50 (acute toxic dose causing the death of 50 % of the test population)
LLNA	Local lymph node assay
L(O)AEL	Lowest (observed) adverse effect level
MAD	Mutual acceptance of data
MIE	Molecular initiating event
MoA	Mode of action
mRNA,	Messenger ribonucleic acid
MPPD	Multiple-path particle dosimetry
MTD	Maximum tolerated dose
NAM	New approach methodology
N(O)AEL	No (observed) adverse effect level
NRC	(US) National Research Council
NTP	National Toxicology Program, US Department of Health and Human Services
OECD	Organisation for Economic Cooperation and Development
OHT	OECD harmonised template
PBT	Persistent, bioaccumulative and toxic
PBTG	Performance-based test guideline
PBTK	Physiologically-based toxicokinetic
PNEC	Predicted no-effect concentration
(Q)SAR	(Quantitative) Structure-activity relationship
QSPR	Ouantitative structure-property relationship
QIVIVE	Quantitative *in vitro* to *in vivo* extrapolation
RAAF	Read-across assessment framework
RhCE	Reconstructed human cornea-like epithelium
RHE	Reconstructed human epidermis
RNA	Ribonucleic acid
SciRAP	Science in risk assessment and policy (tool)
STOT RE	Specific target organ toxicity – repeated exposure
STOT SE	Specific target organ toxicity – single exposure
STS	Sequential testing strategy
SVHC	Substance of very high concern
TER	Transcutaneous electrical resistance
TDAR	T-cell-dependent antibody response
TG	Test guideline
TGR	Transgenic rodent
TMR	Test Methods Regulation
TTC	Threshold of toxicological concern
UVCB	Substance of unknown or variable composition, complex reaction products or biological materials
vPvB	Very persistent and very bioaccumulative
WoE	Weight-of-evidence

List of terms

3Rs principle	Principle of "Replacement, Reduction and Refinement" of animal use, which ultimately lead to the development of non-animal approaches. It was first defined by the scientists William Russell and Rex Burch in "The Principles of Humane Experimental Technique"(1959). (See below for specific definitions for each term.)
Accuracy	The closeness of agreement between test method results and accepted reference values. It is a measure of test method performance and one aspect of relevance. The term is often used interchangeably with "concordance" to mean the proportion of correct outcomes of a test method.
Adaptation	Refers to specific and general adaptation rules under the REACH Regulation and the BPR, which allow fulfilling the required standard information requirement by omitting, replacing by other information or adaptation by other way. There may be also conditions that require further information. The duty holders must state the adaptation and give the reasons for each adaptation.
Adverse outcome pathway	An adverse outcome pathway (AOP) describes a logical sequence of causally linked events at different levels of biological organisation, which follows exposure to a chemical and leads to an adverse health effect in humans or animals. AOP is endpoint-specific and thus substance-agnostic. Similar approaches can be used for mechanisms which may not necessarily lead to an adverse outcome (such as endocrine modes of action).
Benchmark dose (BMD)	The dose corresponding to a specified small increase in effect over the background level.
Benchmark dose lower confidence limit (BMDL)	The lower 95 % confidence interval of a benchmark dose.
(Substance) Category	Group of substances with physicochemical, toxicological and ecotoxicological properties that are likely to be similar or follow a regular pattern as a result of structural similarity.
Defined approach	A defined approach to testing and assessment consists of a fixed data interpretation procedure (DIP) used to interpret data generated with a defined set of information sources, which can be used on its own, or together with other information sources within an IATA, to satisfy a specific regulatory need.
Endpoint	Observable or measurable inherent property/data point of a chemical substance. It may refer to a physicochemical property (e.g. vapour pressure), to degradability, or to a biological effect that a given substance has on human health or the environment (e.g. carcinogenicity, irritation, or aquatic toxicity).
Ex vivo test	An *ex vivo* test is conducted outside the living organism using

	tissues or organs obtained from an animal.
Hazard	The potential for an adverse health or ecological effect. The adverse effect is manifested only if there is an exposure of sufficient level (OECD GD 34).
High-content method (HCM)	High-content method refers to a high-content screening (HCS) or a high-content analysis (HCA) that describes a set of analytical methods using automated microscopy, multi-parameter image processing and visualization tools to extract quantitative data from cell populations.
High-throughput screening (HTS)	A method which involves an automated-operation platform, data processing and control software to quickly conduct many biochemical, genetic or pharmacological tests. It usually allows to test a large number of substances at the same time.
Integrated approach to testing and assessment (IATA)	A structured approach which strategically integrates and weighs all relevant data to inform on a potential hazard and/or risk and/or the need for further targeted testing. It leads towards minimising vertebrate animal testing used for hazard identification (potential), hazard characterisation (potency) and/or safety assessment (potential/potency and exposure) of a chemical or group of chemicals.
In chemico test	Abiotic assay that measures chemical reactivity or other physicochemical properties of substances.
Integrated testing strategy (ITS)	Integrates different types of (eco)toxicological data and information as part of a decision tree to consider the need for further testing to assess the safety of a substance.
Intrinsic properties	In the context of this report, the intrinsic properties of a substance are considered to be purely related to the hazards of the substance, without taking into account (internal or external) exposure. Thus, information on absorption and, for example, on releases from materials into artificial fluids mimicking biological fluids, is related to exposure and may affect the toxicity but not the intrinsic properties of a substance.
In vitro test	Literally stands for "in glass" or "in tube". The test takes place outside the "body" of an organism, usually involving isolated organs, tissues, cells or biochemical systems.
In vivo test	Test conducted within a living organism.
In silico test	Computer-based methods e.g. (Q)SARs. These may be called "non-testing information".
Lowest observed adverse effect level (LOAEL)	The lowest dose level of a test substance in a given test that causes an observed and significant adverse effect on the test species when compared with the controls.
New approach methodologies (NAMs)	This concept has been used as an umbrella for various approaches utilising non-animal methods and technologies which may also allow multiple investigations from a high

	number of samples at the same time. In this report we present the approaches individually and do not generally use this collective term.
Non-animal approaches	Non-animal approaches in this report include all approaches which do not involve **new** animal testing. Thus, in addition to *in vitro* approaches, this term covers approaches which use existing information from animal studies but do not require new animal tests. *In vitro* approaches are included although they may need serum or cells from animals, but they are not actual animal testing. The definition encompasses the use of individual non-animal approaches, such as *in vitro* methods and QSARs; the use of combined and stepwise approaches, such as integrated testing strategies (ITS); or integrated approaches to testing and assessment (IATA) (including defined approaches). In the context of the REACH Regulation, non-animal approaches relate to the use of *in vitro* and *in silico* methods, grouping and read-across (Article 13(1)): "*Information on intrinsic properties of substances may be generated by means other than tests, provided that the conditions set out in Annex XI are met. Information shall be generated whenever possible by means other than vertebrate animal tests, through the use of alternative methods, for example, in vitro methods or qualitative or quantitative structure-activity relationship models or from information from structurally related substances (grouping or read-across)*."
No observed adverse effect level (NOAEL)	The highest dose level of a test substance in a given test that does not cause any observed and statistically significant adverse effect on the test species when compared with the controls.
PBTK model (physiologically–based toxicokinetic model)	PBTK models provide simulated concentration versus time profiles of a substance and its metabolites in plasma or an organ of interest and simultaneously allow for estimation of maximum plasma concentrations, absorption kinetics, distribution kinetics, and elimination. Physiologically-based pharmacokinetic (PBPK), physiologically-based biokinetic (PBBK) and physiologically-based kinetic (PBK) models are used as synonyms.
Performance standards	Standards, based on a validated test method, that provide a basis for evaluating the comparability of a proposed test method that is mechanistically and functionally similar. Included are: (i) essential test method components; (ii) a minimum list of reference substances selected from among the substances used to demonstrate the acceptable performance of the validated test method; and (iii) the comparable levels of accuracy and reliability, based on what was obtained for the validated test method, that the proposed test method should demonstrate when evaluated using the minimum list of reference substances.
Prediction model	A formula or algorithm (e.g. rule or set of rules) used to convert the results generated by one or more tests into a prediction of

	the (toxic) effect of interest. Also referred to as decision criteria.
Quantitative structure-activity relationship (QSAR) and structure-activity relationship (SAR)	Theoretical (mathematical) models that can be used to predict in a quantitative or qualitative manner the physicochemical, biological (e.g. (eco)toxicological) and environmental fate properties of a substance from the knowledge of its structure and properties. A SAR is a qualitative relationship linking a (sub)structure to the presence or absence of a property or activity of interest. A QSAR is a quantitative (regression) model that relates a set of predictor variables (substance structure) to the potency of the response variability (property or activity of interest).
(Grouping and) Read-across	Approach using information from a structurally analogous substance, or a substance with similar mechanisms/modes of action, to predict the properties of a substance. This approach can be used under the REACH and CLP Regulations and the BPR. For the purposes of the REACH Regulation (Article 13(1)), (grouping and) read-across is considered by ECHA to be a replacement (if available information can be used) or reduction (if new animal studies are proposed but less than without read-across) method. Similarly, under BPR, read-across is an adaptation approach where information from an structurally analogous substance may be used to predict the properties of a substance. Information from other substances, structurally similar or with similar mechanisms/modes of action, can be used under the CLP Regulation.
Reduction	Any approach that will result in fewer animals being used to achieve the same objective. It includes maximising the information obtained per animal as well as reducing the number of animals used in the original procedure and/or limiting or avoiding the subsequent use of additional animals.
Refinement	Modification of any procedures or husbandry and care practices from the time the experimental animal is born until its death, so as to minimise the pain, suffering and distress experienced by the animal and enhance its wellbeing.
Relevance	Description of the relationship of the test to the effect of interest and whether it is meaningful and useful for a particular purpose. It is the extent to which the test correctly measures or predicts the biological effect of interest. Relevance incorporates consideration of the accuracy (concordance) of a test method.
Replacement	Any methods, strategies or approaches which enable to avoid the use of live animals to achieve the same objective.
Sensitivity	The proportion of all positive results/active substances that are correctly classified by the (non-animal) test. It is a measure of accuracy for a test method that produces categorical results, and is an important consideration in assessing the validity of a (new, non-animal) test method.
Specificity	The proportion of all negative results/inactive substances that are correctly classified by the (non-animal) test. It is a measure

	of accuracy for a test method that produces categorical results, and is an important consideration in assessing the validity of a (new, non-animal) test method.
Standardised test methods	Test methods validated and standardised at the international level (OECD, ECVAM).
Test (or assay)	Experimental system set up to obtain information on the intrinsic properties or adverse effects of a substance.
Test battery	Group of tests used for one specific purpose, e.g. several individual non-animal approaches, which are combined to cover one *in vivo* endpoint.
Validated test method	A test method which has successfully passed the validation studies process, aiming at determining the relevance (including accuracy) and reliability for a specific purpose. To gain regulatory acceptance, validated test methods also have to show that they are fit for purpose.
Validation	Scientifically-based process by which the reliability and relevance of a method are established for a specific purpose.
Vertebrate animal	Animal that belongs to the subphylum *Vertebrata*; animal with a backbone or spinal column.
Weight-of-evidence (WoE) approach	Approach which can generally be described as a stepwise process/approach of collecting all available evidence, evaluating their quality, integrating and weighing them, to reach a conclusion on a particular property with a (pre)defined degree of confidence. The WoE approach normally requires expert judgement. This is a general approach for evaluating (eco)toxicological properties when information from various sources is available and may allow to follow the weight-of-evidence (WoE) adaptation rule (see below). The approach is also used in evaluating if the properties meet the CLP criteria.
Weight-of-evidence (WoE) adaptation	One of the general adaptation rules under the REACH Regulation and the BPR where information from several independent sources may allow assumption/conclusion of a particular hazardous property of a substance. Results from non-animal approaches may be used as pieces of evidence in a WoE adaptation, usually together with other types of information. However, there may be sufficient weight-of-evidence (WoE) from the use of newly developed test methods, not yet included in the test methods referred to in Article 13(3), or from an international test method recognised by the Commission or the Agency as being equivalent, leading to a conclusion that a substance has or has not a particular dangerous property.

Introduction

Scope

This report reviews the current status of and the near and medium-term future expectations for the regulatory relevance and acceptance of non-animal approaches for information requirements involving vertebrate animals for registration under the REACH Regulation and for applications of biocidal active substance approval under the BPR. This also includes consideration of using the study results to apply the criteria for classification under the CLP Regulation (hazard classification) and for risk assessment (e.g. DNEL and PNEC setting).[2]

The provisions of using vertebrate animal testing as a last resort is included in the REACH Regulation and the BPR and is also reflected in the CLP Regulation. All these regulations promote the use of non-animal approaches according to the 3Rs principle of replacement, reduction and refinement of animal use for testing. Therefore, duty holders have to consider all possibilities to use non-animal approaches that would fulfil information requirements and hazard classification: only if it is not possible to gather sufficient reliable data through non-animal approaches should animal testing be conducted. The present report focuses on approaches that can be used to reliably investigate potential (toxic) effects and bioaccumulation properties relevant for human health and the environment while avoiding the use of vertebrate animals as far as possible.

This report considers the challenges in using non-animal approaches and considers what actions could be taken to overcome them. For instance, integrated approaches to testing and assessments (IATAs) may pose a particular challenge because they cannot be validated for regulatory purposes in the conventional manner, except when defined approaches are used. IATAs allow the use of flexible approaches and WoE approaches using expert judgement, which are more difficult to standardise and are more complex to evaluate compared to single test methods.

This report does not replace any ECHA Guidance Documents and is not primarily addressed to duty holders. This report aims to communicate the opportunities and limitations of using non-animal approaches in a regulatory context involving a wide audience – including registrants under the REACH Regulation, applicants under the BPR, Member State competent authorities, NGOs, the research community and other stakeholders – to understand how these methods may be used.

Where to find information on updated and new test methods?

The main source of information on how to fulfil the information requirements under the REACH Regulation and BPR is the ECHA Guidance Documents. ECHA also frequently updates its website on updated and new test methods (https://www.echa.europa.eu/support/oecd-eu-test-guidelines). With this web page, ECHA supports duty holders in understanding how these methods may be used to meet the information requirements. For example, the role of new test guidelines within testing strategies is described, when appropriate. This information is provided before ECHA's guidance is formally updated.

Use of non-animal approaches

Non-animal approaches and methods (including evidence from humans) may be used under the REACH and CLP regulations and the BPR in different ways, such as:

- to fulfil the information requirements under the REACH Regulation and the BPR:

[2] See list of relevant legislation in Appendix 2 to this document.

  - when the non-animal approach is an information requirement;
  - to adapt or support adaptation of a standard information requirement (e.g. read-across or WoE);
  - to trigger a need of further information to address a concern (REACH or BPR):
    - to address further a particular concern under the REACH Regulation or the BPR;
    - to investigate specific mechanisms or modes of action under the BPR.
  - to support an information requirement.
- to define a hazard classification under the CLP Regulation as (one of the) main or supporting information depending on CLP criteria;
- to support a proposal for inclusion of substances of very high concern (SVHCs) in the Candidate List;
- as information sources in risk assessment reports.

The applicability of non-animal approaches and the related uncertainties and total confidence of the conclusion may differ and depends on the regulatory context and purpose. At the time of adoption of the REACH Regulation in 2006, the time horizon for replacing animal testing for systemic endpoints could not be estimated. It was however predicted that, for example, methods for skin sensitisation were seven to nine years away [1]. Indeed, validated non-animal methods for skin sensitisation have been recently incorporated into the REACH Regulation.

Non-animal approaches are a valuable tool for the evaluation of the available information, although they may not always provide sufficient evidence when considered alone.

Promotion of non-animal approaches

Promotion of non-animal approaches is among the objectives of the REACH, CLP and Biocidal Products regulations. However, this objective shall not undermine the main objectives of REACH – to ensure a high level of protection of human health and the environment.

The application of non-animal approaches in REACH registrations by registrants is analysed by ECHA in a tri-annual report on the use of alternatives to testing on animals, in accordance with Article 117 (3) of the REACH Regulation [2].

ECHA stays committed to promote the development and use of non-animal approaches. In this regard, ECHA closely follows the scientific developments in the field and collaborates internationally. Currently this includes the best practices to use the information from non-animal approaches and contribution to the discussion on how the information from non-animal approaches could be applied in the context of hazard classification. ECHA also promotes the familiarity of the newly-developed methods and approaches among its own staff, scientific committees and stakeholders and regularly updates both its guidance and its website to include information on updated and new test methods to support the duty holders in how these methods may be used under different regulations. Furthermore, ECHA gives proactive scientific and technical advice to the European Commission on both generic aspects related to non-animal approaches and on details of the methods/approaches. ECHA also supports the European Commission in updating the specific annexes.

Test methods, information requirements and adaptations

Non-animal validated test methods are normally included first into the OECD test guidelines (TGs), and only after that are they included in the EU Test Methods Regulation (EC) No 440/2008 (TMR), although this is not a formal requirement. The TMR provides the test protocols (usually based on existing OECD test guidelines) in all EU languages. However, amending the TMR with new methods is a time-consuming process and relevant OECD TGs and potentially other internationally accepted test methods may be considered being appropriate

according to REACH Article 13(3) prior to their inclusion in the TMR. This is reflected e.g. in the information requirements in REACH Annexes VII-X or in ECHA guidance, as appropriate. When regarded as applicable and relevant for the given regulatory purpose, they are normally included in the formal information requirements of REACH and/or BPR. Authorities normally recognise the new relevant method and inform duty holders of its applicability even before the formal incorporation of a new OECD TG to the EU regulations. Furthermore, even if a new method has not yet been accepted by the Commission and/or ECHA to fulfil the information requirement, or has not yet been taken up in *Regulation 440/2008* and the relevant REACH Annexes VII–X, it may nevertheless be used in a WoE adaptation approach in accordance with REACH Annex XI, Section 1.2. WoE adaptation allows the use of not yet internationally accepted non-animal approaches.

The information requirements under the REACH Regulation and the BPR differ mainly by the tonnage-based approach applied for REACH and the "core dataset" and "additional dataset" approach for the BPR. Under both regulations, non-animal approaches are mostly applied under "adaptation" rules. However, for certain endpoints, non-animal approaches are already specified as (part of) an information requirement itself, in which case they are mandatory, "default" requirements.

The legislation includes both adaptation rules specific to an information requirement and general adaptations rules that can apply to all information needs. These rules are described in the REACH and BPR Annexes, and ECHA has elaborated upon them in the REACH Guidance on Information Requirements and Chemical Safety Assessment (IR&CSA), Guidance on BPR Volumes III (Human Health) and IV (Environment), and in the Practical Guide "*How to use alternatives to animal testing to fulfil your information requirements for REACH registration*".

Non-animal approaches in CLP Regulation

Regarding classification obligations, Article 7(1) of the CLP Regulation states that "*where new tests are carried out for the purposes of this Regulation, tests on animals within the meaning of Directive 86/609/EEC* [replaced by Directive 2010/63] *shall be undertaken only where no other alternatives, which provide adequate reliability and quality of data, are possible*". Article 8 of the CLP Regulation also refers to all other means of generating information, including the general adaptation rules provided in Section 1 of Annex XI to the REACH Regulation (i.e. existing data, WoE, QSAR, *in vitro* methods, grouping of substances and read-across approach). If new studies are to be conducted, they must follow the test methods described in the EU TMR, other international methods recognised by the European Commission or ECHA as being appropriate, or follow internationally recognised sound scientific principles, and be validated according to international procedures. In its preamble, the CLP Regulation highlights that the EURL ECVAM plays an important role in the scientific assessment and validation of non-animal approaches. In addition, these methods shall be carried out in accordance with good laboratory practice (GLP) or other international standards recognised as being equivalent by the Commission or the Agency, and with the provisions of Directive 86/609/EEC, if applicable. It should be noted that there is no obligation to perform tests for the purpose of CLP only. Therefore, all available data may be used for the classification of a substance under the CLP Regulation. If REACH or BPR information requirements have been fulfilled for an endpoint, the corresponding data should also be adequate for classification (including sub-categorisation, where applicable) for the corresponding hazard class. This also means that to be considered compliant with the information requirements, the data need to be applicable for classification (and risk assessment).

It should be noted that, although there are already legal provisions for the use of non-animal data under the CLP Regulation, in practice their use for classification is limited and the CLP Regulation does not explicitly mention criteria for the results from *in vitro* tests for all health hazards endpoints. However, CLP allows for a WoE determination using expert judgement decisions on classification, and also non-animal data is regularly relied on in those cases.

Structure of the report

General considerations on all the above aspects of use of non-animal approaches for regulatory purposes under the REACH, CLP and Biocidal Products regulations are presented in the first part of this report (Part A), while the second part of the report (Part B) addresses the information requirements for each endpoint separately. In Part B, each section gives an overview of the currently available non-animal approaches that can be used to minimise animal testing, as well of the challenges to their use and future perspectives. Information requirements, CLP criteria, relevant test methods and specific adaptation rules per endpoint are considered in Appendix 3 to this document.

Part A: General considerations

A.1 Non-animal approaches

A.1.1 Background

Non-animal approaches in this report include all approaches which do not involve **new** *in vivo* testing. Thus, in addition to *in vitro* approaches, this term covers approaches which use existing information from animal studies but do not require new animal tests. For instance, QSAR model predictions or grouping (categories or read-across) adaptations are almost all based on existing *in vivo* test data.

Before conducting any animal study, all the available information (physicochemical properties, grouping, (Q)SARs and expert systems, *in vitro* data, human data and animal data) needs to be assessed. Further guidance is available in that respect within the Guidance on the Application of the CLP Criteria, Guidance on IR&CSA (Chapters R.6, R.7a, R.7b and R.7c), and Guidance on BPR (Volumes III and IV). Most published studies focus on demonstrating the presence of (eco)toxic effects, while the publication of results reliably and adequately showing no (eco)toxicity would also help avoid unnecessary investigations and animal testing.

The REACH Regulation highlights in particular the possibility to use *in vitro* methods, (Q)SARs, grouping and read-across (Article 13(1)) to avoid animal testing: "*Information on intrinsic properties of substances may be generated by means other than tests, provided that the conditions set out in Annex XI are met. In particular for human toxicity, information shall be generated whenever possible by means other than vertebrate animal tests, through the use of alternative methods, for example, in vitro methods or qualitative or quantitative structure-activity relationship models or from information from structurally related substances (grouping or read-across)*." Thus, the REACH Regulation refers to information which is generated by means other than vertebrate animal tests and uses the term "alternative methods". For the purpose of this report, we use the term "non-animal approaches" to cover all the means other than vertebrate animal tests for the substance. Any combination of non-animal approaches, which may also include information from observations in humans and information from existing animal studies, may also be used under a WoE adaptation (Annex XI, Section 1.2), as this allows for bringing together several independent sources of information. An IATA developed for a particular purpose is one such approach, which can integrate information from many different sources such as *in vivo* and *in vitro* methods, QSAR predictions, adverse outcome pathways (AOPs), etc.

To fulfil the aim of ensuring that the use of a substance and mixtures containing the substance is safe for human health and the environment, the potential hazardous properties of the substance need to be identified and investigated. The methods used for such investigations must produce relevant and reliable results adequate for both risk assessment and classification (including sub-categorisation) and labelling. This concept applies to both whether the information requirements are fulfilled with the results from the studies described in information requirements or with data used to adapt these information requirements for example from non-animal approaches. Indeed, under the REACH Regulation (column 2 of Annexes VII-X or Annex XI) and the BPR (column 3 of Annexes II or Annex IV), the stated information requirements (e.g. for information from a reproductive toxicity study) may be adapted, i.e. replaced by other information that allows equivalent regulatory conclusions or outcomes on the hazardous properties and risks of a substance to be reached with a similar level of confidence. Non-animal approaches which are not yet explicitly included in the standard information requirements of the REACH Regulation and the BPR must be used in the context of adaptations. In many cases, this means their use as supporting evidence in grouping and read-across or within a WoE adaptation.

Although not specifically mentioned in the REACH Regulation and the BPR, the 3Rs principle underpins the provisions related to animal testing. The 3Rs refer to: the **replacement** of an animal test with a test that uses non-animal systems, invertebrate species or early-stage vertebrates, or other non-animal approaches; the **reduction** of the number of animals used in a test; and the **refinement** of a test to enhance animal wellbeing and selecting methods causing as little pain or distress for the animals as possible. Refinement may also be achieved through methods using lower vertebrate species (for example fish instead of mammals). This principle was developed by Russell and Burch in 1959 [3] and, based on Directive 2010/63/EU on the protection of animals used for scientific purposes and other specific regulations referring to animal testing, there is an obligation to consider the 3Rs in all safety evaluations (see also the European Commission web pages on 3Rs: http://ec.europa.eu/environment/chemicals/lab_animals/3r/alternative_en.htm.)

More specifically, the **replacement** approach has been defined by Russell and Burch as "*any scientific method employing non-sentient material which may in the history of animal experimentation replace methods which use conscious living vertebrates*" (as cited by Balls in 1994 [4]). They distinguished between **relative replacement**, in which animals would still be required but would not be exposed to any distress in the actual experiment, and **absolute replacement**, in which animals would not be required at any stage. Relative replacement includes, for example, killing animals to collect cells or organs for conducting *in vitro* tests replacing the need for animal tests. For the purpose of this report, these two replacement approaches are considered as non-animal approaches as they do not involve new *in vivo* testing.

Reduction of the number of animals used may be achieved, for example, by using the so-called limit tests, where only one (high) dose level is used instead of several dose levels for a substance with expected low toxicity. The use of a limit dose is mentioned in most standard *in vivo* test guidelines for toxicity testing. Another possibility may be to combine the investigations of two different toxicity endpoints into one study. This may be possible for certain information requirements, such as mutagenicity and repeated-dose toxicity (see the endpoint-specific sections in Part B). Approaches that integrate *in vitro* and *in vivo* methods in a stepwise manner (such as those concerning serious eye damage/eye irritation, and skin corrosion/irritation) may also be considered as reduction methods, as new testing is not conducted systematically but is dependent on the results from the previous step. Including the measurement of new parameters, for example, to investigate endocrine activity, in existing test methods is also a relevant approach and may allow to obtain more information from the same number of test animals.

Usually, there are no study- or endpoint-specific **refinement** possibilities recommended in the internationally accepted *in vivo* test methods, since test conditions are considered as having been already optimised at the OECD and/or EU level. However, general methods to improve housing, husbandry and care as well as refined experimental procedures exist and should be used. Repeating invasive operations or sampling should be avoided in new studies if existing data is reliable and adequate.

It is foreseen that less invasive methods – such as imaging techniques (e.g. magnetic imaging resonance) and investigations using smaller sample sizes – will be developed in the future, thereby reducing potential pain and distress of test animals. Obtaining more information from the same sample may also reduce the number of animals needed, which could be considered as both a refinement and a reduction.

The European Commission's Joint Research Centre (JRC) operates a database service on non-animal approaches (DB-ALM). The DB-ALM is a public, factual database service that provides evaluated information on development and applications of advanced and non-animal approaches in biomedical sciences and toxicology, both in research and for regulatory purposes (see https://ecvam-dbalm.jrc.ec.europa.eu/ and the annual status reports on https://eurl-

ecvam.jrc.ec.europa.eu/eurl-ecvam-status-reports).

In the following sections, various non-animal approaches and some important aspects that should be taken into account when using non-animal approaches are described, including ways of combining the data for regulatory adaptations.

A.1.2 Concepts, individual methods and techniques

A.1.2.1 Adverse outcome pathway and mode of action

An adverse outcome pathway (AOP) is defined by the OECD as "*an analytical construct that describes a sequential chain of causally linked events at different levels of biological organisation that lead to an adverse health or ecotoxicological effect. AOPs are the central element of a toxicological knowledge framework being built to support chemical risk assessment based on mechanistic reasoning*." [5].

An AOP typically starts with a molecular initiating event (MIE), which is the interaction of the substance of interest with its biological targets, e.g. cellular proteins. The MIE triggers a sequential chain of key events (KEs), which are alterations of biological processes each causing a certain downstream effect. This chain eventually leads to an adverse outcome at the tissue/organ level (like liver fibrosis) or at the organism or even population level (for environmental effects).

It is generally assumed that a certain threshold has to be met to provoke the MIE and each of the following KEs. Thus, a certain concentration of the substance is needed to lead to an adverse outcome. It should also be considered that several upstream KEs might lead to the same downstream event. In that sense, there is not a unique AOP for each adverse health or (eco)toxicological effect; rather, it can be through several, sometimes inter-linked, AOPs that substances can cause a specific adverse effect. AOPs do not include aspects related to the toxicokinetics of a substance [5]. Therefore, additional information, in particular on metabolism, is essential to understand whether, for instance, it is the substance itself and/or its metabolites that trigger the MIE.

Modes of action (MoAs) refer to (eco)toxicological pathways leading to effects. MoAs may not describe mechanisms of action but refer to pathways at a more general level. MoAs and AOPs are different frameworks, although there are similarities. As described above, AOPs concern non-substance-specific biological pathways and the final outcome of an AOP is an adverse effect [5]. By contrast, MoAs are substance-specific and include elements such as toxicokinetics and metabolism, and adversity is not needed to define a MoA. Information on MoAs may be used to support read-across within a group (category) if, for example, it can properly be shown that similar substances have the same MoA. This can be used for both hazard identification (classification) and risk assessment if quantitative information is available.

The knowledge of various potential MoAs and AOPs is limited, thus restricting the development of AOP-based methods and approaches. It is not possible to develop an AOP without knowledge of the MIE and KEs and how they are linked together and lead to an adverse outcome. In addition, the classification criteria set some challenges.

Both toxicological and ecotoxicological AOPs have been developed, so far for some simple endpoints. AOPs help structure knowledge of biological and (eco)toxicological processes and can promote the development of AOP-based non-animal approaches addressing the different biological pathways. This is the case for the first AOP accepted by OECD on skin sensitisation, for which the recommended integrated approach to testing and assessment (IATA) is based on *in vitro*/*in chemico* test methods specific for each of the main KEs (see Section B.4).

Generally, a major limitation to the development of AOPs for (eco)toxicity is the current lack of

necessary knowledge about both the underlying mechanisms and MoAs. Extensive efforts are ongoing, for example, in organ toxicity to close this gap in the future.

For complex (eco)toxicity endpoints, it is challenging to develop AOPs, because many of them actually cover a wide range of mechanisms. This is the case, for instance, for reproductive toxicity, which can relate to both parental toxicity and toxicity to the developing organism (see Section B.8). Although complex endpoints could in principle be split into several processes which are easier to model, there are still many unknown aspects and many biological events that may not be covered. This makes the prediction using these non-animal approaches currently impossible for the complex endpoints: they cannot yet provide the data to conclude on no hazardous/adverse effects for risk assessment and hazard classification categorisation.

The comparison of the results from an AOP-based non-animal approach with the CLP criteria may pose a particular challenge, at the same time as quantitative information would be needed for risk assessment. Results from AOP-based non-animal approaches may show adverse effects adequate for classification, including sub-categorisation, and for risk assessment. If the adverse effects detected are adequate for meeting the classification criteria of the most severe category, then it is clear that any other investigation for that property cannot provide more information for hazard classification. These adverse effects, whether meeting the classification criteria or not, may be used together with additional supporting data for grouping and read-across or within a WoE adaptation, or they may trigger the conduct of a definitive study.

To conclude, AOPs have been developed for some simple endpoints, but not yet for more complex endpoints, and it will require significant further effort to reach the set targets. The real challenge is not the development of individual AOPs, but rather the development of a comprehensive set of AOPs that would cover all possible mechanisms that can lead to adverse effects for a given endpoint. This is particularly important if reliable predictions of the absence of hazard/adverse effects are to be achieved with AOP-based methods or approaches.

For further information
Further details can be found on the OECD web page on Adverse Outcome Pathways, molecular Screening and Toxicogenomics. The first five AOPs were published in the new OECD AOP series The collaborative AOPWiki collects the AOPs under development and provides a free online AOP training course. An example of AOP-based toxicodynamic model and a review of this area have been published [6, 7].

A.1.2.2 *In silico* methods

In silico refers to computational methods such as those based on quantitative or qualitative structure-activity relationships ((Q)SARs), expert systems or physiologically-based toxicokinetic (PBTK) modelling. *In silico* methods do not require new animal studies and are therefore sometimes described as non-testing methods. It should however be noted that the data they are built on and use for making predictions of toxicological effects mainly comes from existing information, including *in vivo* studies. Several *in silico* tools for (eco)toxicity are available and have been extensively characterised in the scientific literature [8, 9, 10] (for details on PBTK modelling, see Section B.1 on toxicokinetics).

QSARs are models that can be used to predict in a quantitative manner, yielding continuous or categorical results, physicochemical, (eco)toxicological, or environmental fate properties of substances based on their structures. In the development of QSAR models, substance structure/molecular descriptor information and experimental information on the endpoint of interest for these substances are collected and used as the basis to derive relationships. Once established, these relationships can be applied to predict the same property for a given query molecule if it is well represented by the training set of substances used to develop the QSAR model (i.e. the prediction for the query substance is within the applicability domain of the

QSAR model).

Expert systems are rule-based with the rules being derived by expert knowledge and/or statistical induction. Some expert systems also incorporate (Q)SAR models and read-across assumptions. These systems typically predict test results based on aggregated data from multiple sources, and may therefore provide a high-level assessment of (eco)toxic potential.

Some QSAR models are trained (developed) to give categorical/binary results: the substance is predicted to have or not to have a particular toxicological property, like mutagenicity, in an Ames test. Such models could be of use in screening and (de)prioritisation. However, ruling out toxicity on the basis that no alert was found within a single SAR system is not normally possible because there may be other mechanisms not included in the model. Therefore, the regulatory use of these (Q)SAR models is challenging: they may detect alerts if models predict effects, but may not if used alone allow to rule out toxicity if they predict no effects.

Another common limitation is that dose-response information required for risk assessment, for example a N(L)OAEL, cannot be derived from current QSAR models. However, for ecotoxicological assessment, many QSAR models have been developed to predict short-term aquatic ecotoxicity, especially in fish, algae and daphnids, and they typically predict LC_{50} or EC_{50} values, which may be used for both hazard classification and environmental risk assessment. For bioaccumulation evaluation, too, some QSAR models can predict actual bioconcentration factor (BCF) values.

The general rules for adaptation mentioned in Section 1.3 of REACH Annex XI and BPR Annex IV state that: "*Results obtained from valid qualitative or quantitative structure-activity relationship models may indicate the presence or absence of a certain dangerous property.*" Thus, the results obtained from reliable (Q)SAR models may allow to adapt standard information requirements and thus reduce testing on living organisms. Results from a QSAR model can be used if: (i) the scientific "validity" (i.e. reliability and relevance) of the model has been established (for instance, following the OECD QSAR validation principles); (ii) the substance falls within the applicability domain of the model; (iii) the prediction is fit for regulatory purpose (i.e. for hazard classification categorisation and risk assessment); and (iv) the applied method is well documented.

Results from QSAR models may also be used as one piece of evidence in a WoE adaptation (Section 1.2 of REACH Annex XI and BPR Annex IV), as supporting evidence in combination with other data when reading across (eco)toxicological properties in a grouping and read-across adaptation (Section 1.5 of REACH Annex XI and BPR Annex IV) and classification and labelling, or within an integrated approach like an IATA.

It should be noted that for complex toxicological endpoints, QSAR prediction models are not currently considered reliable. For instance, reproductive toxicity includes both information from developmental toxicity, sexual function and fertility, and from key parameters which are currently not covered by QSAR (or any other non-animal approach) to the extent that either of them or both information requirements would be fulfilled, for prenatal developmental toxicity (OECD TG 414) and extended one-generation reproductive toxicity study (EOGRTS, OECD TG 443). Therefore, a negative result from a QSAR prediction alone will not be acceptable for complex toxicological properties unless there is additional supporting evidence so that the all elements for that (eco)toxicological property are adequately covered.

To conclude, depending on the predicted property and the regulatory purpose, valid and adequate QSAR data may suffice on their own or may need to be supported by other types of evidence to come to a conclusion. These data can indicate a potential hazard and trigger further testing.

For further information

There are several commercial and public domain *in silico* tools available. For instance, the OECD QSAR Toolbox is free to download and is maintained and developed systematically as software. It can be used to help group substances for read-across based on structural similarity, mechanism of action and metabolism. Another freely available tool is the Database, which provides prediction profiles concerning physical-chemical, environmental fate, ecotoxicity and toxicological (including some endocrine disruption-related) endpoints. Further freely available QSAR/expert systems include the rule-based Toxtree [8], EPISUITE, VEGA, T.E.S.T., Metaprint-2-D-calc., SPARC, etc. JRC QSAR Model Database can be found on JRC's website.

Recommendations on how to apply QSAR can be found in Chapter R.6 of the REACH Guidance on IR&CSA. In ECHA's updated Practical Guide "*How to use and report QSARs*", issues of applicability domain and reporting are included. Advice on pitfalls and good practice in using QSARs are recorded annually in the ECHA Evaluation reports.

A.1.2.3 *In vitro* and *in chemico* methods

In vitro methods usually involve isolated organs, tissues, cells, or biochemical systems. Most common *in vitro* methods are based on cell or tissue cultures consisting of similar cells (monoculture systems) or combining different cell types (co-culture system). Some test methods are based on the culture small organs *in vitro*. These organ cultures are also called *ex vivo* cultures because the organs or part of them are taken directly from exposed or non-exposed animals or humans and cultured under *in vitro* conditions. The isolated chicken eye test method (EU B.48/OECD TG 438) and bovine corneal opacity and permeability test method (EU B.47/OECD TG 437) are examples of such *ex vivo* test methods.

In chemico assays are abiotic assays that measure chemical reactivity. A JRC report explains the basis of the *in chemico* approach to toxicity prediction and reviews the studies that have developed the concept and its practical application since the 1930s, with special attention to studies aimed at the development of QSAR models and grouping and read-across [11]. The main applications identified are related to the assessment of skin sensitisation, aquatic toxicity and hepatotoxicity.

The cells used in *in vitro* methods may be obtained from animals or humans directly (primary culture) or from specifically-developed immortal cell lines usually derived from cancer cells. Alternative techniques of cell immortalisation have been developed and are commercially available, like "conditionally immortalised" cell lines, which can undergo differentiation [12] and have some metabolic competence. Furthermore, embryonic stem cells or other pluripotent stem cells can be used and differentiated into organ-specific types of cells to study the effects of test substances in various tissues [13, 14, 15]. Various stem cell applications are under development [16, 17], such as the *in vitro* neural embryonic stem cell test as a replacement for the neurodevelopmental toxicity test [18, 19].

Cells may be cultured in monolayer, in two layers, or in more complex structures. The most advanced methods aim to mimic the functions of an organ, as inlike 3D skin models or organs-on-a-chip. The organs-on-a-chip microsystems are artificial mini-organs built as multi-channel 3D microfluidic translucent cell culture devices. These organs-on-a-chip may simulate physiological activities, such as the breathing motions in the lung or the peristaltic movements in the intestine, and can be used to study the functions and responses of an entire organ or even multi-organ systems when several microsystems are connected with one another by vascular channels (forming a so-called "human body-on-a-chip"). The methodology is undergoing remarkable research activity biomedical engineering and there are different models of organs-on-a-chip available. So far, organ models implemented in microfluidic devices include the heart, lung, kidney, artery, bone, cartilage, and skin. These systems are expected

to become a promising tool in drug development and future toxicity testing and in providing relevant information to predict toxicity and chemical exposure within the human body at different life stages (e.g. [20, 21]). One limitation of the organs-on-a-chip methodology is the size of the organs, i.e. the low number of cells mimicking each organ. It should however be noted that the different models vary significantly in terms of technology, design and approach and they still need further development, optimisation, standardisation and validation before clear recommendations on their use for regulatory purposes can be made. Currently, results from organs-on-a-chip microsystems may support grouping and read-across and a WoE adaptation as one piece of evidence. However, the information is limited on the organs investigated and cannot show widely "no repeated dose toxicity" or "no specific target organ toxicity".

The combination of high-throughput and high-content techniques with *in vitro* methods allows the use of multi-analyses and multi-sample test systems compatible with automation and reduced amounts of test substance. Not all *in vitro* test methods are compatible with such techniques, which usually not only imply technical and methodological adaptations compared to the classic method (e.g. change of equipment, culture plate formats, culture conditions, read out and analytical approach), but also require a separate validation of the modified test method. For instance, while widely used for screening, the miniaturised versions of the Ames test for mutagenicity testing have not been universally accepted as replacements for standard regulatory testing, although they are described in OECD TG 471. An OECD review of existing data for a possible retrospective validation of the miniaturised Ames tests is ongoing (see Project 4.109 of the OECD Work plan for the Test Guidelines Programme 2017 [22]).

It should be noted that some *in vitro* systems may require the use of animals, like those using primary cells. In addition, serum obtained from animals may be needed for cell culture. Some *in vitro* test methods also require the use of a metabolic activation system (S9 fraction), which is usually obtained from the liver homogenate of rodents treated with a chemical. However, there are efforts ongoing to limit animal use and develop, for instance, serum-free or human cell-based *in vitro* methods.

Currently, *in vitro* testing is an information requirement for skin corrosion/irritation, serious eye damage/eye irritation and mutagenicity under the REACH Regulation and the BPR, and for skin sensitisation under the REACH Regulation. A number of standard *in vitro, ex vivo* and *in chemico* methods for these endpoints are available in the OECD test guidelines programme and in the EU TMR (see Sections B.3, B.4 and B.6, respectively) and should be used to fulfil the information requirements.

Results obtained from suitable *in vitro* methods, i.e. *in vitro* methods developed according to internationally agreed criteria (see Section A.2.3), may indicate the presence of a certain hazardous property or may be important for mechanistic understanding. Such results normally require confirmatory testing unless the conditions for adaptation stated in Section 1.4 of REACH Annex XI and BPR Annex IV are met. However, results from one *in vitro* test alone may not suffice for the adaptation to fulfil the information requirements. Results from several *in vitro* test methods may be used in combination to support a WoE adaptation (Section 1.2 of REACH Annex XI and BPR Annex IV). Additionally, *in vitro* data can be used together with other information as supporting (mechanistic) information for grouping and read-across and as elements within a WoE adaptation (for example, in integrated approaches such as IATAs) for more complex information requirements (Sections 1.5 and 1.2 of REACH Annex XI and BPR Annex IV).

It should be noted that *in vitro* models can only predict effects that are known and for which these models were designed for – *in vitro* models cannot be used to detect unknown mechanisms. Ideally, results from *in vitro* studies should be extrapolated to reflect the situation *in vivo* or in humans to allow considerations of hazardous exposure levels. One of the reasons why so far *in vitro* tests have been successful mostly for qualitative risk assessment

lies in the fact that there is still limited knowledge on (quantitative) *in vitro* to *in vivo* extrapolation ((Q)IVIVE). The (Q)IVIVE approach [23] uses PBPK modelling [24] and provides a useful tool to correlate the *in vitro* concentration-response curves to equivalent *in vivo* or human dose-response relationships and helps determine threshold exposure values for risk assessment (see also Section B.1). However, as described below, several aspects should be taken into account when interpreting the results from *in vitro* test methods and extrapolating them to the *in vivo* situation.

One challenge is that **cell lines may give a different response compared to cells in the body ("wild type cells")** because of their different background and, furthermore, because cultured cells may show alterations in their biology. The *in vitro* environment differs from the *in vivo* environment with respect to oxygen level, medium composition, extracellular matrix composition and the functional interactions between tissues. A better understanding of these differences, the identification of relevant cellular characteristics to help select an appropriate cell line and/or better interpret test results (as for example recommended for some standard mammalian *in vitro* mutagenicity test methods [25]), and the improvement of cell culture methods and approaches (for instance, to better model cell organisation and interactions in the body) may provide some answers to this challenge.

Moreover, cultured cells may lack metabolic capability or have an unbalanced metabolism. This is often compensated for in *in vitro* tests by the addition of an external source of metabolic enzymes (S9 fraction), although this implies using animal samples. Metabolic competence is however not always crucial, notably if the substance is not metabolised or if the focus of the study is on certain local effects.

Depending on the endpoint, consideration of the relevance of the *in vitro* test results includes the requirement to **reflect systemic interactions** of cells, tissues and organs, which is often lacking. Furthermore, most models do not simulate the effects in the whole organism, which is always more complex than a combination of information from several *in vitro* tests. As described above and in Section A.1.2.3, some methods using complex cell culture systems, such as 3D skin models and organs and body-on-a-chip devices, are able to address this aspect to some extent.

The design of *in vitro* methods for complex toxicity properties, such as reproductive toxicity, is especially challenging due to a large number of potential targets/mechanisms associated with this broad area of toxicity. Replacement of complex endpoints by *in vitro* methods does require systems and strategies for adequate reflection of metabolism and systemic tissue interactions. Also, species-specific aspects need to be considered. While many modern assays incorporate aspects of metabolism and biotransformation, they are often still subject to limitations regarding tissue interactions or long-term stability.

In vitro biokinetics inform on the fate of the substance in the *in vitro* test system, including binding to proteins, cell/culture medium partitioning, binding to plastics, etc. This information is critical for the establishment of concentration-response curves and interpretation of the test results because the free concentration and the intracellular concentration of a substance do not directly correspond to the quantity applied. Measuring biokinetics in *in vitro* studies to improve predictivity is essential and allows a better (Q)IVIVE.

The above considerations can be seen as limiting the relevance of information obtained from *in vitro* tests. However, such information may still adequately reflect the toxicity when the lack of networks and function of the whole organisms is not critical. This is the case in particular for endpoints like skin sensitisation. Among the non-animal approaches, *in vitro* methods are often the ones most fit-for-purpose for certain endpoints, whether as stand-alone information (e.g. for irritation and mutagenicity testing) or, for example, as supporting mechanistic information in defined approaches (see Section A.1.3.1 below), IATAs or AOP-based strategies (see Section A.1.2.1 above).

In conclusion, *in vitro* methods can be used as stand-alone information, potentially as part of defined approaches, or as one element in WoE, depending on the purpose and information requirement in question. They may provide important information on mechanisms and MoAs, also support grouping, and read-across. Significant efforts are ongoing to develop *in vitro* methods based on AOPs and organs-on-a-chip technology and apply the QIVIVE approach for *in vitro* to *in vivo* extrapolation of effects.

For further information
Recommendations on the use of the recently adopted *in vitro* methods for each endpoint can be found in Chapter R.7 of the Guidance on IR&CSA, the respective EU test methods or OECD TGs, and ECHA's web page "Testing methods and alternatives".

A.1.2.4 Methods based on "-*omics*"

Various methods are based on "-*omics*", which are large-scale analytical techniques that can be used to support and understand biological and (eco)toxicological mechanisms.

Methods based on "-*omics*" are used to group particular sets of biological molecules produced in cells and provide profiles or "fingerprints" that reflect cell response under different conditions, like after exposure to a substance. These methods may detect a cellular response but whether this response is linked to adversity or an adaptive response, how it relates to the substance dose/concentration, or whether changes are reversible or not, may not be predicted from the results. However, it may be possible to identify a "tipping point", i.e. a substance dose or concentration above which the effect is no longer reversible but leads to an adverse outcome.

The analysis can be done on samples taken from animals and *in vitro* tests. Depending on the focus of the study, different sets of analyses can be done. For example:

- The genome is the genetic material of an organism. It is defined by deoxyribonucleic acid (DNA) sequences divided into chromosomes. The genome is composed of both the genes (coding regions of the DNA) and non-coding DNA. The corresponding "-*omics*" investigation is called genomics.
- The transcriptome is the set of all the messenger ribonucleic acid (mRNA) molecules in one cell or in a specified population of cells. It reflects gene expression at a given moment and may vary depending on environmental conditions. The corresponding "-*omics*" investigation is called transcriptomics.
- The proteome is the entire set of proteins expressed by a genome (at a given time under defined conditions). The levels of mRNAs are not directly proportional to the expression level of the proteins they code for. The corresponding "-*omics*" investigation is called proteomics.
- The metabolome refers to the entire set of small molecules (metabolites) found in an organism, tissue or cell. It may include endogenous metabolites as well as exogenous substances. The corresponding "-*omics*" investigation is called metabolomics.
- The epigenome refers to specific epigenetic modifications involving changes to the chromatin structure, regulating gene expression through, for instance, DNA (de)methylation or histone (de)acetylation. The corresponding "-*omics*" investigation is called epigenomics.
- The regulome refers to transcription factors and other molecules involved in the regulation of gene expression. The corresponding "-*omics*" investigation is called regulomics.

Functional genomics is a field of investigation that focuses on gene (and protein) functions and interactions. It covers function-related aspects of the genome itself, such as mutations, polymorphisms, and molecular activities through the use of different techniques, such as transcriptomics, proteomics and metabolomics. Thus, functional genomics focuses on the

dynamic aspects of cell response – including gene transcription, translation, regulation of gene expression and protein-protein interactions over time (such as during an organism's development) or space (such as in the different body regions) – as well as on studies of natural or experimental functional disruptions affecting genes, chromosomes, RNAs or proteins.

Analyses based on "-*omics*" can be combined with *in vivo* studies or *in vitro* tests to enhance the information obtained from test animals or cells, both in terms of quantity – as they have high data content – and quality – as more insight into the underlying mechanisms can be gained.

Non-animal approaches based on "-*omics*", or test methods integrating these techniques, are still under research and development. The execution of such analyses and the interpretation of their results are yet to be standardised. Therefore, further development is needed before they can be integrated into regulatory decision making. However, "-*omics*" analyses are expected to be able to support grouping and read-across and WoE approaches (under the REACH, CLP and Biocidal Products regulations) and in principle could already be applied for these purposes if the analyses reflect the (eco)toxicity to the organ, tissue or biological system in question.

In conclusion, various "-*omics*" may enhance mechanistic understanding and provide biological profiles at different levels of investigation, and they can also provide a tool for screening and (de)prioritisation for further regulatory action. These methods may also help identify potential biomarkers of exposure to a substance or some biological effects. They do not provide direct information on adversity but may indicate potential toxicity pathways that may lead to adverse health effects.

Analysis based on "-*omics*", approaches integrating such a technique, and the interpretation of related results are still under development and approaches have not yet been standardised for regulatory use.

A.1.2.5 Concluding remarks

A number of non-animal approaches are available. The AOP-based concept significantly increases the possibilities for constructing non-animal approaches that utilise information on biological pathways and disturbances which relate to adverse effects. Analyses based on "-*omics*" are rapidly evolving to provide increasingly accurate information on pathways involved in toxicity and may provide a general view of various molecular and potential functional changes in cells caused by a substance. Organ-on-a-chip technology offers a promising approach to predict organ toxicity and can also reflect interactions between various organs. Many of these approaches are under development, to be used either as stand-alone methods for certain endpoints or as supporting evidence, for example, for read-across. It is important to develop standardised rules for conducting such studies and for how to interpret and present the results for regulatory use.

Together with *in vitro, in chemico, in silico* and other techniques and non-animal approaches, "-*omics*" analysis are also called by the term *new approach methodologies* (NAMs), especially if data are collected using high-throughput screening (HTS) methods or high-content methods (HCMs). HTS and "-*omics*" methods are not toxicity tests in their own right, but rather represent analytical techniques that enhance toxicity tests *in vivo* or *in vitro* by adding new endpoints or parameters, pattern recognition, high-throughput, etc. These methods may be used in screening and (de)prioritisation, in suggesting a mode of action (MoA), in identifying endpoints that can be used in AOP development (or "-*omics*" to interpret read-outs for HTS), as supporting information for grouping and read-across, in integrated approaches such as IATAs, and as elements within WoE determination and adaptations.

The potential of non-animal approaches in a number of regulatory settings was recognised

recently in the context of the REACH Regulation and other international chemical legislation during the ECHA Topical Scientific Workshop "New Approach Methodologies in Regulatory Science", which took place on 19-20 April 2016. A number of key recommendations for the further applications of non-animal approaches in a regulatory context were made (see the corresponding workshop proceedings [26]).

Non-animal approaches are able to confer classical, whole organism-based toxicology with mechanistic information. Non-animal approaches are considered to have great potential in risk-based approaches as described in [27] and [28]. In general, toxic effects could be better understood with the help of changes observed in "-*omics*" information. ECHA will further explore the use of these approaches for priority setting and as supporting evidence in hazard and risk assessment.

In the current situation, non-animal approaches cannot predict the outcome of an animal study, and as such cannot be seen as direct replacements for animal studies in the short term. They do however give information on the effects of chemicals on the organism or parts of it. This information could potentially be used to set safe values of exposure. Thes concepts can be used for (de)prioritisation of chemicals, but could also – at least in principle – be an integral element of a chemicals management system. Further exploration of what levels of protection these approaches could offer, especially compared to current animal study-based systems, is needed. Especially the matters of "over-protection" or "over-regulation of chemicals demand careful consideration in this context.

A.1.3 Methods and approaches to use, integrate and weigh available information

A.1.3.1 Sequential testing, IATA and defined approaches

Different approaches exist to characterise hazards based on the combination of the available data and/or the generation of new data to fill in data gaps. Depending on how these approaches are structured and defined, different terms are commonly used.

The general term **integrated approach to testing and assessment** (**IATA)** is used for a pragmatic, science-based approach of hazard identification and characterisation. It relies on an integrated analysis of existing information coupled with the generation of new information where needed using testing strategies [29] (see also http://www.oecd.org/chemicalsafety/risk-assessment/iata-integrated-approaches-to-testing-and-assessment.htm). An IATA consists of modules or components which are each based on individual information sources. These are grouped together according to, for example, the type of information provided (i.e. *in vitro* data, *in vivo* data, physicochemical properties, etc.) or the "mechanistic level" of the information (i.e. toxicokinetic data, or toxicodynamic data related to KEs within an AOP). An IATA necessarily includes a degree of **expert judgement** in weighing the available information.

Gathering existing information and use of information from non-animal approaches in an IATA may be similar regardless of the decision context. However, generation of new test data (e.g. data from *in chemico* or *in vitro* methods) may differ depending on the scope of the IATA and the evidence collected. Generation of new data should be tailored to reduce the uncertainty of the conclusion. Evaluation of existing information or generation of additional data within an IATA can be performed on the basis of a non-formal WoE approach or by using predefined, structured approaches such as **defined approaches** or their combination as described below.

As defined by OECD [29], "*a defined approach to testing and assessment consists of a fixed **data interpretation procedure (DIP)** used to interpret data generated with a defined set of information sources, that can either be used on its own, or together with other information sources within an IATA, to satisfy a specific regulatory need*". Defined approaches to testing

and assessment are standardised (i.e. rule-based) and may be used as components in an IATA [29]. In a defined approach, data generated by non-animal and animal methods deemed to be fit-for-purpose are evaluated by means of a fixed DIP. The output of a DIP is typically a prediction of a biological effect of interest. A DIP is rule-based in the sense that it is based, for example, on a formula or an algorithm (e.g. decision criteria, or a rule or set of rules) that does not involve WoE determination[3] using expert judgement [29, 30]. The use of defined datasets and a more objective assessment than WoE determination using expert judgement are expected to facilitate the regulatory application of these defined approaches. While no defined approaches have been formally approved yet, in April 2017 the OECD approved a project aiming to analyse the predictivity of several preliminary defined approaches for skin sensitisation. If found acceptable, these defined approaches can be used under the REACH Regulation. These defined approaches are potentially useful for REACH registrants, since they would enable more robust use of the *in vitro* methods specified under the REACH Regulation.

A **sequential testing strategy (STS)** is a fixed stepwise approach for obtaining and assessing test data, involving interim decision steps which, depending on the test results obtained, can be used on their own to make a prediction or to decide on the need to progress to subsequent steps. At each step, information from a single source/method is typically used by applying a prediction model associated with that source/method [29].

An **integrated testing strategy (ITS)** is a term previously used for an approach in which multiple sources of data or information are assessed at the same time by applying a variety of specific methodologies to convert inputs from the different information sources into a prediction. For this purpose, a variety of specific methodologies can be applied, such as statistical and mathematical models [29]. It is to be noted that the concepts of IATA and ITS are still under development [31], and that IATA and ITS have often been used as synonyms. However, as indicated above, an ITS can be considered as a limited type of IATA that uses identified methods, and an IATA has a slightly broader scope than an ITS with respect to the "assessment" part, since it uses a WoE assessment (see NTP). For a proper evaluation of an IATA, the various components and information sources used within the IATA should be well characterised and documented in terms of their applicability, limitations and performance. To this end, several templates exist, for example, from OECD for reporting defined approaches to testing and assessment [29], individual information sources [29], QSARs [32], grouping and read-across strategies [33], and non-guideline test methods [34].

Several examples of IATAs or defined approaches have been published or are under discussion by the OECD for skin corrosion/irritation [35], serious eye damage/eye irritation [36], and skin sensitisation [37], and other IATAs are being developed for non-genotoxic carcinogenicity. A specific strategy for skin sensitisation assessment under the REACH Regulation has been developed based on the above OECD Guidance Documents (GDs) (see Section B.4 on sensitisation and Section R.7.3 of the Guidance on IR&CSA – Chapter R.7a).

The use of an IATA under the REACH Regulation and the BPR falls in most cases under WoE adaptations, unless the components of the IATA strictly correspond to the (standard) information requirements for a property. This may be the case for example for skin corrosion/irritation and serious eye damage/eye irritation endpoints, skin sensitisation endpoints, and mutagenicity endpoints (see Sections B.3, B.4 and B.6, respectively, and the corresponding Sections R.7.2, R.7.3 and R.7.7, respectively, of the Guidance on IR&CSA – Chapter R.7a).

For those endpoints for which *in vivo* testing is still the default information requirement under the REACH Regulation, an IATA can nevertheless be used to structure a WoE adaptation. Such

[3] WoE determination using expert judgement is the terminology used under CLP Regulation for a WoE approach (see Section A.1.3.2).

a WoE adaptation would also be needed for using an IATA to fulfil the information requirements for skin sensitisation under the BPR, since those requirements do not yet reflect the *in vitro/in chemico – in vivo* stepwise approach described in the revised REACH Annex VII. Even if an IATA contains only *in vitro* studies or only QSAR predictions, it should be used under the WoE adaptation as long as there are several independent sources of information used together. It could be possible to build up an IATA for a source substance, and read-across the result to a target substance and use it under a WoE adaptation; however, there are no examples yet for that kind of an approach and its feasibility. In all cases, care should be taken that the IATA addresses relevant aspects of the information requirement and that the substance falls within the applicability domain of the methods used within the IATA.

A.1.3.2 Weight-of-evidence approach and weight-of-evidence adaptation

In a **weight-of-evidence (WoE) approach**, all relevant information should be collected together, which may then suffice to allow for a conclusion to be made without further studies. The WoE approach is a normal way of considering all information available for an information requirement or (hazard) assessment under the REACH, CLP and Biocidal Products regulations.

There is no single formal definition of WoE. The principles are described in Section 1.2 of REACH Annex XI and BPR Annex IV and within ECHA Guidance Documents (Guidance on IR&CSA - Chapters R.2 to R.4). Further practical aspects of the use and reporting of WoE under the REACH Regulation can be found in Section 4.1 of ECHA Practical Guide "*How to use alternatives to animal testing to fulfil your information requirements for REACH registration*". A number of descriptions of WoE approaches are available from WHO/IPCS, SCENHIR [38], EFSA [39], US OSHA [40] and OECD (see for example [35]).

A WoE approach can be defined as follows: "*WoE can be generally described as a (stepwise) process/approach of collecting evidence and weighing them to reach a conclusion on a particular problem formulation with (pre)defined degree of confidence*."

WoE approaches may be used for different purposes, e.g.:

- to justify a concern and trigger further information needs;
- to reach a regulatory conclusion such as hazard classification; or
- to reduce animal testing by using different lines of evidence to adapt a new animal test with already existing information (i.e. to fulfil the information requirement).

The WoE approach can be divided into several steps:

1. Problem formulation
2. Collection of information
3. Assessment of quality of individual evidence
4. Integration and assessment of overall evidence
5. Confidence levels and remaining uncertainty
6. Conclusion

First, in the problem formulation, the regulatory target of the WoE approach should be defined. After defining the purpose of the WoE, pieces of information should be collected and assessed individually for their quality. This is an important step because information should not be disregarded based on its poor quality – rather, it should still be presented but its weight (value) can be lowered down, even to zero, if necessary. This is due to transparency reasons, to show that all the information has been considered. If there are several pieces of information addressing the same aspect (e.g. QSAR model predictions and results from *in vitro* methods addressing a particular mechanism or mode of action (MoA) for developmental toxicity), these pieces of information may be grouped together as one line of evidence for this mechanism or MoA, and the consistency of this information has to be considered.

During integration and assessment of the overall evidence, each piece of evidence is weighed (valued) against all other pieces of evidence and a value for overall evidence is considered and may be quantified (e.g. using confidence levels). Consistency of the evidence needs to be evaluated and the quality of individual pieces of evidence is considered in the weighing. Plausibility (credibility) and human relevance should also be considered. At this stage, the completeness of the evidence is also taken into account, meaning that it should be evaluated if there are investigations missing which could critically affect the outcome of the WoE analysis regarding the information requirement.

A WoE approach (and WoE determination as referred to in CLP, see below) normally requires expert judgement. To make this WoE approach using expert judgement transparent and comprehensible, it is essential that all information used, its assessment and the conclusions drawn are fully documented and justified.

The purpose of the WoE approach could be, for example, to assess if a substance's toxicological property needs to be classified and categorised or not under the CLP Regulation (**WoE determination**). The CLP Regulation requires the following for hazard determination:

"*Where the criteria cannot be applied directly to available identified information, or where only the information referred to in Article 6(5) is available, the weight-of-evidence determination using expert judgement shall be applied in accordance with Article 9(3) or 9(4) respectively*"; and "*A weight of evidence determination means that all available information bearing on the determination of hazard is considered together, such as the results of suitable in vitro tests, relevant animal data, information from the application of the category approach (grouping, read-across), (Q)SAR results, human experience such as occupational data and data from accident databases, epidemiological and clinical studies and well-documented case reports and observations. The quality and consistency of the data shall be given appropriate weight. Information on substances or mixtures related to the substance or mixture being classified shall be considered as appropriate, as well as site of action and mechanism or mode of action study results. Both positive and negative results shall be assembled together in a single weight of evidence determination.*"

When WoE is used to replace (adapt) the need of a standard test according to Section 1.2 of REACH Annex XI or BPR Annex IV, a WoE adaptation needs to be substance- and case-specific as well as hazard-based, and needs to address the relevant (standard) information requirement that is adapted (i.e. it is endpoint-specific). A WoE adaptation is specific to an information requirement (endpoint), and the elements for which information should be available depends on information requirements. For instance, for repeated-dose toxicity, it is important that the target organs have been reliably identified and that the level (severity) of organ toxicity has been investigated. In addition, adequate and reliable documentation with a rationale or justification needs to be provided (by the duty holders). Specific considerations of the information requirement are presented at relevant endpoint parts below. In such a WoE adaptation, **independent sources of information** are collected which may lead to the scientifically justifiable assumption/conclusion that a substance has or does not have a particular hazardous property. The assessor can conclude on the basis of the WoE adaptation that no further information is required for a specific information requirement and can conclude on a particular hazardous property of a substance.

All the key parameters or elements of an information requirement should be addressed within a WoE adaptation. The level of the evidence (completeness) needed is dependent on various aspects, for example, observed effects, their nature and severity, and the (eco)toxicological profile of the substance. The use of a valid WoE adaptation for a given standard information requirement (as specified in REACH Annexes VI-X or BPR Annex II) and the use of the actual standard information should be expected to lead to a similar regulatory outcome. In REACH evaluation processes, ECHA assesses the use of WoE as an adaptation of information

requirements (as specified in REACH Annex XI, 1.2) proposed by the registrant. Under the BPR, the applicant may use Annex IV, 1.2, subject to the reporting Member State's approval. The pieces of evidence are evaluated individually and together with respect to being adequate and relevant to meet a specific information requirement.

A WoE adaptation aims to inform on intrinsic hazardous properties of a substance and is not described as a risk-based approach in the information requirements in the REACH Regulation or the BPR. Therefore, exposure-related justifications cannot be used as elements within a WoE adaptation. Results from new test methods may also be considered under this adaptation as pieces of evidence among other available information or as stand-alone information from an international test method that is recognised by the Commission or the Agency as being equivalent.

The Guidance on IR&CSA – Chapters R.2 to R.6 and the Endpoint-Specific Guidance Documents R.7a, R.7b, R.7c provide the background information on the interpretation and application of the WoE evidence as phrased in REACH Annex XI of and BPR Annex IV. For information on the quality of data, see A.2.4 below.

A.1.4 Grouping and read-across adaptation

To **read-across** means to predict (instead of to measure) a property or test outcome for one substance (the target substance) on the basis of a property or test outcome obtained with another substance (the source substance) by following certain rules. **Grouping of substances** refers to the formation based on certain rules of a category/group of substances whose physicochemical, toxicological and ecotoxicological properties are likely to be similar or follow a regular pattern, usually as a result of structural similarity. Application of the group concept requires that physicochemical properties, human health effects and environmental effects or environmental fate may be predicted from data for reference substances within the group by interpolation to other substances in the group (read-across approach). This avoids the need to test every substance for every endpoint. If certain pieces of information on the properties of category/group members are missing, it may be possible to apply read-across to fill data gaps. Grouping of substances and read-across is one of the most commonly used adaptations for meeting information requirements in registrations submitted under the REACH Regulation [2] and the BPR. If the grouping and read-across is applied correctly, experimental testing can be reduced, as there is no need to test every substance in a group.

Essential elements as per Annex XI of REACH or BPR Annex IV
The prerequisite for applying the read-across technique is that there is a scientifically sound read-across hypothesis (i.e. a justification) and data are available to support it. Such justification must explain why and how read-across can be applied for the property under consideration. The principles of read-across described below are based on the REACH Regulation, but the same principles apply for the BPR. If the registrant proposes a read-across which is reliable, its acceptability can be fully evaluated only when the data, either existing or obtained through the generation of new information, are available. It therefore follows that a read-across relying on toxicokinetics considerations would require such information to be available.

Read-across can be used between two substances or it can be applied to a group of substances. In both cases, the substances need to be structurally similar so that it can be assumed that physicochemical, toxicological and ecotoxicological properties are similar or form a regular pattern. Within such groups, data gap filling is considered by taking the study result(s) for one or more source substances and using them for another "similar" substance (target substance), taking into account not only the numerical result of the test but also the whole set of effects (including hazard classification) established with the source substances. There has to be at least one study with a source substance which is relevant and of good quality to serve the purpose of data gap filling. However, as a general rule, more rather than

fewer source substances with data are needed to make more reliable predictions, unless the read-across can be based on (bio)transformation into common substances.

As an adaptation in REACH Annex XI or BPR Annex IV, the results should: (i) be adequate for classification and labelling and/or risk assessment; (ii) have adequate and reliable coverage of key parameters addressed in the corresponding test method; (iii) cover an exposure duration comparable or longer than the corresponding test method; and (iv) be supported by adequate and reliable documentation. It is important to note the requirement that information should be adequate also for classification and labelling.

Read-across under the CLP Regulation
Using read-across under the CLP Regulation to apply classification does not necessarily follow exactly the same rules as read-across for REACH or BPR information requirements. CLP refers to "substances chemically related" to the substance under study and "group entr", whereas under the REACH Regulation and the BPR more detailed requirements are described. For read-across under the CLP Regulation, the identifiable functional groups and a powerful mode of action (MoA) acting across a substance group may already allow the same hazard classification (e.g. anticoagulants and specific metal substances), and a structural similarity is less critical in a WoE determination. Where a toxic MoA is well described for chemical groups (as with e.g. pyrethroids, organophosphates and metals), the toxic properties of one known substance can be inferred for a similar unknown substance. Although complex, it is currently possible to use read-across for hazard classification when all necessary considerations are taken into account. However, this becomes increasingly difficult in cases where the MoA is not understood. This is typically the case for substances that are not biocides, such as industrial chemicals, or that have little human health data available.

Grouping and QSAR toolbox
While searching for analogue substances and formulating a hypothesis, (Q)SAR applications may be useful. (Q)SAR programs can be used to group substances based on their structures for potential read-across (see Section A.1.2.2 above). The main ECHA Guidance on how to apply QSAR for grouping and read-across can be found in Chapter R.6 of the Guidance on IR&CSA. This Guidance is closely linked to the OECD Guidance on grouping [33], which was updated in 2014. An illustrative example of a grouping of substances and read-across to support companies in complying with their obligations under the REACH Regulation is available from ECHA's website. An example on how metabolomics could be used to support read-across has been published [41].

Bridging data (supporting data)
For complex endpoints in Annex IX and X, *in vivo* bridging data are usually necessary to support the read-across (see in particular Sections B.5 and B.8 for repeated-dose toxicity and reproductive toxicity, respectively). For example, to read-across information for sub-chronic toxicity (90-day study), information on 28-day repeated-dose toxicity (OECD TG 407 or OECD TG 422 or similar) should in principle be available for both the source and target substances. Such information may help substantiate predictions of similar systemic toxicity potentials for the substances.

Regarding the REACH Annex VIII information requirements, read-across may be challenging in the case of limited availability of supporting *in vivo* data (if information from higher-tier study requirements is not available). Therefore, other supporting information should be collected (although it may not be sufficient on its own), e.g. from *in vitro*, *in silico*, "*-omics*" or toxicokinetic methods. To read-across information for a 28-day repeated-dose toxicity study (OECD TG 407) or a screening study (OECD TGs 421/422), the only supporting *in vivo* information available may be the results from an acute toxicity study. However, such information supports rather poorly predictions for repeated-dose toxicity and information from non-animal approaches and toxicokinetics may provide further support.

Toxicokinetic data are valuable in supporting read-across in all cases, for any information requirement involving systemic exposures *in vivo*, and in particular for studies of repeated-dose and reproductive toxicity. Toxicokinetic data are also generally valuable in supporting read-across regarding metabolism/transformation rates, e.g. for substances forming common substances (toxicants) by (bio)transformation.

Similarly, when predicting environmental effects by grouping and read-across, information on the environmental fate of the target and source substances is essential. The justification should explain why the target and source substances behave similarly in the environment and are equally bioavailable to the organisms during testing (or form a trend based on a property limiting bioavailability). This information may include hydrolysing and degradation properties, physicochemical properties such as water solubility, dissociation, volatility and partitioning constants (organic matter, lipids).

For prediction of environmental effects, bridging data are usually necessary to support the read-across. For instance, to read-across information for long-term toxicity to fish (e.g. OECD TG 210), information on short-term toxicity to fish (e.g. OECD TG 203) should be available for both the source and target substances.

Challenges and limitations of the read-across

Although read-across is a widely used adaptation to avoid unnecessary animal testing, it has some challenges and limitations that need to be considered. First of all, read-across is endpoint-specific. This means that even if the read-across allows prediction for repeated toxicity, further similarity is not automatically plausible for other endpoints, such as reproductive toxicity. There is always uncertainty involved when information from another substance is used to predict hazardous properties of another substance. However, the regulatory outcome should be similar whichever way the information is generated. Read-across can be accepted based on available information, but it is possible that in certain cases new information may challenge the accepted read-across and it needs to be reconsidered. In addition to differences in the main components of a substance, there may be significant (unknown) differences in other components and impurities. For UVBC substances all the components (and the chemical (sub)structures) and their amounts are not known, and also their hazardous properties cannot be known, which is an extra challenge for read-across. To reduce the uncertainties, a worst-case approach is used for read-across (see the RAAF document [42] for details).

Further information

The application of read-across in REACH registrations by registrants is analysed in a tri-annual report on the use of alternatives to testing on animals, in accordance with Article 117 (3) of the REACH Regulation [2]. Advice on pitfalls and good practice in using read-across is recorded on an annual basis in the ECHA Evaluation reports.

Since a wide range of possible scientific explanations may be used for read-across cases, ECHA developed a Read-Across Assessment Framework (RAAF) document, including both human health and environmental endpoints [42]. The RAAF provides a structured framework and principles for the scientific examination of read-across cases, so that the crucial aspects are assessed in a consistent way. The RAAF outlines the main hypothesis on which read-across predictions can be made and builds a number of scenarios around those. The framework also provides a useful basis for establishing the critical elements for assessing the reliability of a read-across. It therefore serves also as a framework to identify and apply new experimental techniques aimed at supporting read-across/grouping. Lessons learned from read-across case studies for repeated dose-toxicity have been published [43]. Specific considerations on the complexity of using read-across for multi-constituent and UVCB (unknown or variable composition, complex reaction products or biological materials) substances are also covered by the RAAF [44].

A.2 Non-animal approaches: focus areas, challenges and outlook

A.2.1 Focus areas for scientific development

Reservations have been expressed regarding approaches relying solely on *in vitro* and other non-animal methods because it has been considered unlikely that non-animal approaches would provide information with the same prediction accuracy as animal studies [45, 46, 47].

Still, some requirements for lower-tier *in vivo* studies have now been replaced by *in vitro* testing. Under the REACH Regulation, this is the case for serious eye damage/eye irritation, skin corrosion/irritation and skin sensitisation. These replacements are important achievements, not only in terms of animal welfare, but also in terms of implementation of new scientific methods in line with current state-of-the-art regulatory science.

However, until now, only a few higher-tier *in vivo* studies have been replaced by *in vitro* testing. For instance, after almost 10 years from the publication of the vision and strategy for toxicity testing in the 21st century [46], the uterotrophic assay, an *in vivo* study, has only recently been replaced by a testing battery of *in vitro* tests to be used for screening substances [48]. Due to the more complex nature of higher-tier information requirements, specific non-animal approaches that could directly replace vertebrate animal tests are not yet available and not foreseen in the near future. Additionally, it is unlikely that one-to-one replacement can be achieved, as most of the currently available test methods are not stand-alone test methods. Rather, a combination of non-animal methods and approaches and their integration into testing and assessment strategies is expected to address some of the difficulties to evaluate complex endpoints.

Factors with a perceived influence on regulatory acceptance and use of non-animal approaches have been explored and discussed with recommendations by Schiffelers *et al*. [49, 50]. Two of the recommendations seem to be highly relevant: **data sharing of both *in vivo* and *in vitro* data to diminish the existing uncertainties of 3R models (industry)** and **the creation of safe harbours for data sharing (regulators)**. Several challenges put forward already in 2009 and referred to by Busquet and Hartung in 2017 [51] include, for example, considerations on testing strategies and threshold setting. There are developments addressing many of these challenges, such as the *in vitro* to *in vivo* extrapolation, but many still need more work. Data sharing is already one essential element to reduce tests on vertebrates under the REACH Regulation (one substance – one registration), however, data sharing is needed worldwide to enhance to use of existing information. The duty holders can use adequate existing data as well as testing strategies such as read-across.

Another example of a challenge to the development and regulatory acceptance of non-animal approaches is provided by Sauer *et al*. [52] in relation to implemented stepwise approaches (skin corrosion/irritation, serious eye/damage/eye irritation and skin sensitisation. In general a WoE (adaptation) needs to be built to replace the concurrent *in vivo* test, as most of the currently available test methods are not stand-alone test methods. A future challenge for OECD's work will be the regulatory acceptance of testing strategies – and mutual acceptance of data resulting from those. For the developed OECD performance-based test guideline (PBTG) for defined approaches and test methods for skin sensitisation, see Section B.4 on skin sensitisation.

Further aspects described below and related, for instance, to the interpretation of results from non-animal approaches, their relevance to humans and their suitability for regulatory use, may explain why the development and acceptance of non-animal approaches for additional endpoints are challenging and how they could be supported.

Mechanistic understanding and technical development

As described in Section A.1.2.3, *in vitro* systems present some differences compared to normal cells regarding behaviour, response, metabolism, cellular organisation and interactions as well as substance biokinetics. These need to be taken into account in the interpretation of results from *in vitro* systems and the comparison of results from the use of various different methods. Some limitation may be easily addressed, such as metabolic capacity by adding metabolic enzymes. Using co-culture allows to address cellular interaction to some extent, and the organs-on-a-chip approach allows in addition investigations of dynamic interactions of several organs, which is a significant improvement to static culture conditions.

Organs- and body-on-a-chips now allow to model *in vitro* functional organs and organ systems whose response to a substance is expected to be physiologically more relevant than that of more simple cell culture systems (see also Section A.1.2.3). They should also enable direct observation of the impact of the substance on the functions of the organs and help identify potential adverse effects and the underlying mechanisms. Development is ongoing to improve the methods so as to allow for their use to predict organ toxicity for those organs that can be included and address dose-response and adversity. The methods still need more development, optimisation, standardisation and validation. The approaches require microfabrication, microelectronics and microfluidics to be able to achieve the required conditions to model physiological responses *in vitro*. Information on biokinetics in *in vitro* studies improve the (quantitative) *in vitro* to *in vivo* extrapolation (Q)IVIVE [23] with the help of PBPK modelling [24].

Among the non-animal approaches, *in vitro* methods are often the ones most fit-for-purpose for certain endpoints, whether as stand-alone information (e.g. for irritation and mutagenicity testing) or, for example, in defined approaches, IATAs or when using an AOP concept. *In vitro* methods may provide information on specific aspects within an AOP concept or an IATA such as cellular effect or receptor binding.

A major difficulty common to higher-tier endpoints is that a variety of processes are involved and toxic substances can act through different MoAs. A more comprehensive **mechanistic understanding of the adverse effects** is therefore needed to develop non-animal test methods and approaches able to model and/or assess them. This understanding is also useful for the design of further *in vivo* studies. Mechanistic data are finally a prerequisite for the translation of the (*in vitro*) effects into meaningful parameters for regulatory use.

Research activities are addressing this matter. For instance, the European project EU-ToxRisk aims to "*drive the required paradigm shift in toxicological testing away from 'black box' animal testing towards a toxicological assessment based on human cell responses and a comprehensive mechanistic understanding of cause-consequence relationships of chemical adverse effects*". The project focuses on repeated-dose systemic toxicity, developmental and reproductive toxicity and integrates advancements in cell biology, "*-omics*" technologies, systems biology and computational modelling with an ultimate goal "*to deliver testing strategies to enable reliable, animal-free hazard and risk assessment of chemicals*". This holistic approach is demanding but supported. It may be challenging to interpret information from some high-content non-animal approaches such as "*-omics*": for example, the expression of various genes may be up-regulated and/or down-regulated differently in different organs and tissues. The nature of the effects, and whether they are showing adaptive responses or reflecting adversity, may be unclear.

The **ongoing work on AOPs** is another way of addressing the complexity of some adverse effects. Improved knowledge of critical KEs and KE relationships is expected to help develop *in vitro* and *in silico* approaches that provide more reliable information on relevant toxicological outcomes. AOP-based test systems may remain "simple" by focusing on some KEs only and without necessarily reproducing complex biological networks in animals as this is the case for the skin sensitisation AOP (see Section B.4). However, in such an approach where the different

KEs of an AOP are covered by different test methods, the complexity may be increased in the interpretation of the results and overall assessment. Moreover, a comprehensive assessment of the adverse effect would imply that all the underlying mechanisms and KEs are covered by an exhaustive AOP network, which may not be (easily) achievable for some complex endpoints.

Reliable evidence of certain MoAs may indicate a potential adverse effect and can be used to support read-across without the need to conduct a new animal study (e.g. cholinesterase inhibition, inhibition of pituitary hormones). Currently, information on a certain MoA cannot yet used to predict an adverse outcome. For example, anti-androgenic properties of a substance do not automatically predict reproductive toxicity and information on a (anti-androgenic or any other) MoA cannot be used alone to classify a substance's hazard as belonging to the most severe category of reproductive toxicity. This is because there are no hazard classes for MoAs and, thus, information on a MoA does not lead to any hazard classification which could provide similar risk management measures to the respective adverse outcome (e.g. reproductive toxicity). However, there is a special relationship between mutagenicity and genotoxic carcinogenicity. Currently, under the REACH Regulation, mutagenicity is already used as a reliable predictor for genotoxic carcinogenicity and hence *in vivo* carcinogenicity studies normally do not need to be conducted for genotoxic substances already classified as germ cell mutagens in category 1B. However, mutagenicity studies alone cannot lead to carcinogenicity classification, but the risk management measures and downstream consequences are the same with the germ cell mutagens and genotoxic carcinogens.

Thus, in the future, reliable predictions of adverse effects may be developed for certain properties, or at least generally to adverse outcomes, but currently those predictions cannot replace observation of adverse effects *in vivo*.

***In vitro to in vivo* extrapolation and extrapolation to humans**

An important aspect determining the acceptability of data from non-animal approaches, in particular *in vitro* methods, for regulatory use is the **translation of results** to *in vivo* and/or human toxicity.

As for the interpretation and extrapolation of *in vivo* data, **inter-species differences** should be considered when analysing *in vitro* test results. While it is fair to assume a high degree of similarity between higher vertebrates, the remaining differences often still introduce uncertainty into the toxicological assessments.

For the extrapolation of results from *in vivo* studies to humans, the uncertainty related to species differences – which are mainly related to toxicokinetic and toxicodynamic differences – is addressed by using an overall inter-species "uncertainty" factor composed of an allometric scaling factor and a substance-specific remaining uncertainty factor (for details see Chapter R.8 of the Guidance on IR&CSA).

For the extrapolation of results from *in vitro* studies, the origin of the cells used (animal or human) may play a role and, in specific cases, it may be useful to do some comparative studies and tests in both animal cells and human cells *in vitro* and compare the results to those from animal studies, to better understand the potential species differences and predict human relevance.

However, also other parameters should be taken into account in the extrapolation: e.g. the type of cells used in the *in vitro* system *vs.* the type of the target cells *in vivo* or in humans; the target species of the extrapolation (animal or human); the dose/concentration-response relationship; and the administration mode *vs.* the route of exposure *in vivo* or in humans. Most of these parameters can be taken into account through PBPK modelling (see also Section B.1)

A systematic evaluation of the information from different species should provide information on

the uncertainties in animal studies as well as *in vitro* approaches. Species concordance and species differences *in vitro* are also addressed at OECD level (update of OECD GD 150).

Information on toxicokinetic *in vitro* can also give indication on the most relevant species to be tested, thus, resulting in reduction of animal testing in the absence of alternatives.

The ability of non-animal approaches to reflect **route-specific aspects** also needs to be considered because the toxicokinetics, target organs and toxicity of a substance may differ depending on the route of exposure (i.e. oral, dermal or inhalation route), except maybe for local effects. If the aim of testing is to identify the intrinsic properties of a substance, then non-animal approaches may adequately identify systemic effects (although route-specific aspects may not be known) but may not necessarily be adequate for hazard classification (for a specific route) and for risk assessment (DNEL derivation).

Another challenge to the translation of information from non-animal approaches were recently highlighted in a position paper by Tralau *et al*. [53]: "*A major reoccuring barrier for alternative in vitro tests, apart from metabolism and tissue interaction, is the translation of molecular or physiological biomarkers into the quantitative parameters required for risk assessment. Targeted research as well as pilot assessments could help to address this*." When testing a substance *in vivo*, the dose/concentration selection should address toxicity by including levels high enough to detect adverse effects and identify the hazards, or reach the limit dose when applicable, to allow hazard classification and the identification of a NOAEL value or calculation of a benchmark dose for DNEL derivation (see Chapters R.8 and R.10 of the Guidance on IR&CSA). To replace or adapt *in vivo* studies, non-animal approaches should be able to provide this type of dose/concentration-response information (and information for the hazard classification from the same or separate non-animal approach). One difficulty is to link the *in vitro* effective concentrations to equivalent doses in animals or humans. The QIVIVE modelling approach has been developed to overcome this issue and is an essential element for *in vitro*-based risk assessment [23] (see also Section B.1). However, work is still needed to implement this approach in regulatory practice and, even if some models show promise for risk assessment, it is not clear at this stage how QIVIVE could be used alone for hazard classification.

Substance-specific aspects

Substance specific aspects should be taken into account. For example, the introductory sections to REACH Annexes VII-X and Point 9 of BPR Annex II point to specific adaptations of the standard information requirements, as *in vivo* testing must be avoided with corrosive substances at concentration/dose levels causing corrosion.

The **nature of a substance** should be taken into account when planning to use non-animal approaches. Even mono-constituents are not presented by one molecular structure, as impurities can have an impact on the (eco)toxicity. Multi-constituent substances consist of more than one main constituent and they usually also have impurities.

UVCBs are substances of unknown or variable composition, complex reaction products or biological materials. Properties of various components may be different or unknown, leading to differences in (eco)toxicology, toxicokinetics and environmental fate, and cannot be fully investigated with the methods due to possible limitations with respect to applicability domain, solubility, unplanned binding, lack of or unbalanced metabolism of the components, etc. Metabolism and abiotic degradation are of the highest importance when applying non-animal approaches, both in determining the relevant toxicant (e.g. as target for read-across) and in deciding whether a certain substance with all its constituent falls in the applicability domain of a non-animal approach. *In vivo* or *in vitro* toxicokinetic information is challenging to obtain for multi-constituent and UVCB substances. Unknown metabolism and abiotic degradation of such substances may hamper the use of non-animal approaches and effective grouping and read-

across. For read-across, supporting data such as "-*omics*" in relevant metabolically competent cell systems (regarding information requirement to be read-across) could be considered. *In vivo* bridging data could also be improved with "-*omics*" data.

An ECHA document on using read-across for multi-constituent and UVCB substances within the RAAF describes key issues in assessing and addressing the complexity of read-across for these substances [44]. Furthermore, it contains example model cases to illustrate this complexity.

Nanomaterials are materials which contain nanosized particles. A nanomaterial is a natural, incidental or manufactured material containing particles in an unbound state, as an aggregate or an agglomerate and where, for 50 % or more of the particles in the number size distribution, one or more external dimensions are in the size range 1-100 nm.
However, there is also list of nanomaterials which have dimensions outside of the 1-100 nm range. The European Commission adopted a recommendation on an overarching definition of a nanomaterial in 2011 (Recommendation on the definition of a nanomaterial (2011/696/EU)).

Nanomaterials are considered to be like any other substance; some may be hazardous, some not. However, nanoforms of substances can be expected to have hazardous properties different from those of the bulk form and should therefore be assessed separately. There may be a need to take into account specific considerations in their assessment such as in toxicokinetics. Existing standard test methods may need modifications for nanomaterials [54], and some have been updated to address the specific features of nanomaterials. Some other tests still need to be updated in that respect, e.g. the solubility test. Some properties of the nanomaterials linked to MoAs can be investigated using predictive methods such as *in vitro* cultures. *In vitro* testing for nanomaterials presents a number of limitations [55], and needs to be taken into account.

JRC has published a report on the availability and applicability of computational methods for nanomaterials, based on the outcome of the Nanocomput project [56]. Specific guidance and recommendations for nanomaterials related to IR&CSA and Registration under the REACH Regulation are also available on the ECHA website. Further guidance has been published by EFSA [57] and SCCS [58].

Performance and standardisation

The high complexity and limitations of standard *in vivo* animal tests, such as low predictivity due to species differences, is used to argue that the **benchmark of performance** to be overcome with non-animal approaches is not very high and this increases the chance for success for non-animal approaches [59]. However, so far there is no evidence to allow the comparison of the benchmarks of performances. It has also been proposed that with the chemical regulations, more precautionary approaches for limit values and classification and labelling may be taken. In this context, the exact prediction of adverse effects and adverse effect levels may not be as important as identifying a likely safe exposure range for the highly variable human population [59, 60]. However, the current regulatory system is not built this way (see the explanation for precautionary principle in Section A.2.2). Ultimately, the main goal is the safe use of substances, and the benchmark of performance of non-animal approaches should meet this goal.

One of the focus area is the **wide variety of different non-animal approaches**. This is conceptually similar to the challenge of agreeing on which of the high variety of potential animal testing approaches are to be used. However, to date, a higher level of harmonisation has been achieved, since many (eco)toxicological endpoints and measured parameters are described and integrated within one test guideline. Conversely, non-animal approaches testing individual endpoints are usually not to be used as stand-alone methods, and therefore harmonisation is more useful for integrating assays for several KEs and KE relationships. The performance of these **defined approaches** with respect to applicability domain, sensitivity and specificity and/or for other relevant features and limitations must be characterised – as

done, for example, in the validation of non-animal approaches – and clearly indicated in the EU/OECD TGs. Until this is achieved, the proper use and interpretation of the results from such defined approaches remains challenging. Furthermore, several non-animal approaches may need to be developped for risk assessment and hazard classification and sub-categorisation as it may not necessarily be possible to have a single approach applicable to both purposes. Therefore, further development is needed on how to apply information from non-animal approaches in a regulatory context.

Suitability of information from non-animal approaches on classification and labelling and risk assessment is discussed more specifically in Section A.2.8.

Summary

In conclusion, there are still various scientific challenges relating to non-animal approaches to adequately provide the information currently required by regulations for systemic effects. However, several initiatives on scientific approaches are addressing these challenges (e.g. AOPs and organs-on-a-chip). Information on toxicokinetics and knowledge on MoAs and AOPs and the connection to human biology will help to support the regulatory acceptance of non-animal approaches as stand-alone information. In general, non-animal approaches may be useful for a number of regulatory uses, and they could be used imaginatively and flexibly, driving changes in regulatory hazard assessment practice in the future.

A.2.2 Addressing uncertainty

For many methods, more scientific developments are needed to increase the confidence that the safety level they provide is sufficient to ensure a high level of protection.

The safety level setting is a policy decision. The level of confidence to be reached for each hazard is reflected by the information requirements of the REACH Regulation and the BPR. Both regulations expect an equal level of safety to be achieved by using the methods, including non-animal approaches, specified as information requirements or as adaptations.

Therefore, new non-animal approaches to be used to fulfil or adapt information requirements should provide the confidence for a safety level at least as high as the one supported by the current information requirements.

However, non-animal approaches will often need to include several methods and may provide a different type of information compared to the standard data requirements. Acceptance of information from non-animal approaches sometimes may mean accepting a kind of uncertainty different from that of standard information required under the REACH Regulation or BPR. Another, but related, issue is that the acceptability of uncertainty is higher if the non-animal approaches indicate that the substance has a property with a value which is far away from a regulatory decision point (such as a classification cut-off) compared to if the value is very close to such a regulatory decision point.

Different *in silico models* (such as QSAR prediction models) for some information requirements may give different prediction results for a substance. In principle, **the predictions made by a given model** should be 100 % reproducible. Lack of trust towards these models may however result from other considerations – poor predictivity, limitations due to the applicability domain, lack of mechanistic relevance, or that the predicted property is not a regulatory endpoint. The predictions are based on the input data, and transparency of the chemical space and algorithms is essential. As an outcome, a numerical value for the prediction may not adequately express the uncertainty (or confidence) involved. Unlike a given QSAR prediction model, the outcome of the same study design from an *in vivo* or *in vitro* study may vary and always includes some uncertainty.

The confidence level (level of uncertainty) is influenced by substance-specific, method-specific

and species-specific aspects, which are sometimes not fully known or understood. The method-specific aspects are most likely to be best understood, such as which parameters are investigated or what the statistical power of the study is. However, there is no international agreement yet on reproducibility, relevance and uncertainty from complexity of the long-term standard animal tests, although work is in progress at OECD level for the field of carcinogenicity [61]. This is expected to provide information on the selection of suitable reference information for the validation of non-animal approaches in relation to carcinogenicity and as well as maybe an objective benchmark for minimal performance of new approaches.

One challenge may relate to the ultimate certainty of applicability and acceptability of using non-animal approaches to fulfil the regulatory requirement from the user side (e.g. registrants/applicants) and uncertainty on the safety level they provide from both the registrants/applicants and authorities' sides. The greatest certainty can only be provided when non-animal approaches are included in information requirements. More confidence in the performance of non-animal approaches is achieved when they are internationally accepted test methods because they have been validated, which means that their performance has been assessed, internationally reviewed and agreed and/or that their performance has been internationally evaluated and recorded. In the acceptance of *in vitro* methods it should be distinguished whether they indicate a certain dangerous property or the results show no effect. In the case of no (eco)toxicity, further steps are required (e.g. confirmation that *in vivo* is not necessary). Thus, the uncertainty is mostly in the cases with no effects.

The concept of precautionary principle covers specific circumstances where scientific evidence is insufficient, inclusive or uncertain but there are indications for reasonable grounds of concern for human health or the environment that necessitate preventive decision-taking. Whether or not to take action on the basis of the precautionary principle is usually decided at the risk management phase of a process. The application of the precautionary approach should not be confused with a prudential approach to address scientific uncertainty in risk assessment. Thus, it cannot be used to cover the increased uncertainty in cases where less robust information is used to fulfil certain information requirements. Precautionary actions, for instance applying a more severe hazard class, cannot be used to improve the confidence level. Under the CLP Regulation, classification and sub-categorisation are based on certain criteria, without a possibility to apply more severe classification based on a precautionary action. Under the REACH Regulation and the BPR, the information requirements must be fulfilled with adequate data. Although an additional assessment factor can be used to cover for poor quality of data in risk assessment, the use of non-animal approaches as such introduces another type of uncertainty that cannot be simply compensated by applying an extra assessment factor. Case-by-case consideration is needed.

It is recommended to include assessment of uncertainties (or confidence) for any data used for adaptations. This would help increase the transparency of documentation and assessment of the potential need for additional assessment factors or further data (see for example [62]). ECHA has published a guidance on uncertainty (see Chapter R.19 of the Guidance on IR&CSA), and EFSA has a revised draft guidance on uncertainty [63]. The possibility to evaluate confidence/uncertainty stemming from information on non-animal approaches would greatly improve their use.

A.2.3 Validation of non-animal approaches

Validation is defined as the process by which the reliability and relevance of a particular approach, method, process or assessment is established for a defined purpose [31]. In particular, validation of a test method means **assessing different performance parameters** to ensure that the method does what it is meant to do in a reproducible and accurate way. Usually, the predictive capacity of the method is characterised as well as its limitations during the validation process. Since the validation process is conducted under standardised and controlled conditions, it is generally required to facilitate and/or accelerate the international

(regulatory) acceptance of non-animal approaches and methods. Following adequate validation studies demonstrating the utility and the applicability of a test method/approach, it may be considered for adoption by regulatory institutions such as the European Commission and the OECD.

As stated in Section 1.4 of REACH Annex XI and BPR Annex IV, a **valid *in vitro* method** can be considered suitable if it is developed according to internationally agreed criteria. In the EU, the EURL ECVAM is in charge of coordinating the validation of alternatives to animal testing and of promoting their scientific and regulatory acceptance.

The EURL ECVAM formal validation process ensures a science-based and independent evaluation of test methods and approaches with the aim to establish their overall performance and fitness for a given purpose. The formal validation includes a four-step procedure:

(1) assessment of test method submissions taking PARERE (EURL ECVAM's Network for Preliminary Assessment of Regulatory Relevance) and ESTAF (ECVAM Stakeholder Forum) opinions into account;
(2) planning and conduct of validation studies, possibly in collaboration with EU-NETVAL (Network of Validation Laboratories) laboratories;
(3) coordination of independent scientific peer review by the EURL ECVAM Scientific Advisory Committee (ESAC);
(4) development of EURL ECVAM recommendations on the validity status of test methods taking stakeholder and ICATM (International Collaboration on Alternative Methods) input into account.

The validity of an individual non-animal approach is assessed by addressing its relevance and reliability for a specific purpose, and includes the characterisation of its applicability domain and limitations. These include sensitivity and specificity considerations, and the method should not under- or overestimate the effect. By specifically addressing questions such as "*which signalling pathways are involved*?", "*do the in vitro methods reflect physiology in vivo*?", "*are there species differences involved*?", "*how do the findings on the cellular level translate to the organ and the organism*?", and "*at what internal doses is adversity observed in vivo and how can this be transferred into in vitro concentrations*?", it will be possible to evaluate the relevance and reliability of data generated in relation to its intended use and hence to consider its regulatory applicability.

The validation process of the OECD follows the same principles and is essential for the approval of OECD TGs (and EU test methods in the TMR). The process is done according to OECD GD 34 [30]. The conduct of an independent validation is important and it can be done by an organisation, not necessarily EURL ECVAM, as long as the principles of OECD GD 34 are followed. The OECD test guidelines programme as well as non-animal approaches that undergo validation available by ECVAM should be regularly consulted for any updates.

Both relevance and reliability are important aspects of validation, and to be acceptable for regulatory purposes, a non-animal approach should thus be both (pre)validated and scientifically valid. In certain cases it may be considered whether methods not yet validated can be used to predict a positive effect, e.g. under a WoE adaptation. However, some non-animal approaches cannot be formally validated following the traditional approach, or the validation process is considered too laborious. In those cases, other practices are followed. For instance, the QSAR methods are characterised in terms of five OECD principles, including applicability domain which should be reported together with the prediction result. For other methods, characterisation based on performance factors could be used. Sometimes the term "validation" can be used in different contexts, i.e. "mechanistic validation" as explained for instance in an EFSA report on developmental neurotoxicity [64].

While the validation of non-animal approaches is usually performed against animal data, the use of human data was possible for the validation of such approaches for skin sensitisation.

The validation of non-animal approaches for the higher-tier endpoints is challenging especially if the validation is required to be done using information gained from humans. Such human data may not be available or suitable for validation and other means then need to be used. The existence of human data in this context is of great advantage but not a requirement, neither in *in vitro* nor *in vivo* contexts.

Whether a test can be validated or not depends on its degree of standardisation, the existence of mechanistic (e.g. AOP) information that allows to put the test into the context of a defined approach (or IATA), and the availability of a sufficiently large set of tests and standard reference substances. Both OECD and ECVAM are currently working on emerging approaches for validation (e.g. concept for defined approaches).

Also IATAs which contain WoE assessment elements need an international regulatory agreement based on scientific principles and this may also be considered to represent a validation. However, this latter approach will use more mechanistic information and less data correlation analysis. Validation is a time- and resource-consuming process and there is no clear solution to how to facilitate and accelerate this step in method development.

In general, a WoE (adaptation) needs to be built for an IATA to replace the concurrent *in vivo* test, as most of the currently available test methods are not stand-alone test methods. A future challenge for OECD's work will be the regulatory acceptance of testing strategies and mutual acceptance of data resulting from them.

Finally, read-across and other case-by-case WoE determination/adaptations may only be based on regulatory agreements, principles and guidance. To gain regulatory acceptance, validated methods also have to show that they are fit for purpose.

A.2.4 Quality

Another important aspect to take into consideration when evaluating the information gathered, including data from non-animal approaches, is data quality.

The evaluation of data quality includes assessment of:

- adequacy of the information for hazard/risk assessment and classification and labelling purposes;
- relevance of the information for a particular hazard identification or risk characterisation; and
- reliability of the information in terms of clarity and plausibility of the findings.

These terms were defined by Klimisch *et al*. [65] (see also [30] and Chapter R.4 of the Guidance on IR&CSA). For the assessment of existing information (non-human data – physicochemical properties, grouping, (Q)SARs and expert systems, *in vitro* data; human data; and animal data) further guidance is available within the Guidance on the Application of the CLP Criteria, the Guidance on IR&CSA – Chapter R.4 and the Guidance on BPR Volume III – Parts B+C *Human Health Assessment & Evaluation*. There are also publications on good *in vitro* method practice (e.g. [66]).

The Guidance on IR&CSA – Chapter R.4 provides the elements for the assessment of data quality (on the basis of reliability, relevance and adequacy) and completeness. Data quality may be evaluated using the Klimisch criteria for the assessment of reliability of guideline-conform data. For other types of evidence (e.g. QSAR, use of read-across, non-standard *in vitro* assays), ECHA Guidance provides criteria for their assessment.

There are also tools such as CRED (**C**riteria for **R**eporting and **E**valuating ecotoxicity **D**ata) [67] and ToxRTool (**Tox**icological data **R**eliability Assessment **Tool**) for *in vivo* and *in vitro* toxicity data, which may be of help in evaluating the inherent quality of data and increase the

transparency of the evaluation. Criteria of the Science in Risk Assessment and Policy (SciRAP) tool are developed to facilitate and increase the use of peer-reviewed *in vivo* toxicity and ecotoxicity literature in regulatory risk assessment of substances. The recent article from the Federal Institute for Risk Assessment (Bundesinstitut für Risikobewertung; BfR) discusses an approach to assess the relevance and reliability of experimental data (*in vivo* and *in vitro*) from guideline-compliant studies as well as from non-guideline studies published in the scientific literature in the specific context of uncertainty and risk assessment of pesticides [68].

Data quality is also linked to the quality of the test method used to generate them. The knowledge of how a study was carried out, and consequently of its relevance and reliability, is a prerequisite for the evaluation of the study results. Where there is more than one study for one endpoint, the greatest weight is given to the studies that are the most relevant and reliable (see also Section R.4.2 of the Guidance on IR&CSA – Chapter R.4). In particular, validation of the method (see Section A.2.3) and GLP status of the study (see also "*OECD Principles on Good Laboratory Practice*" [69] and OECD "*Guidance document for describing non-guideline in vitro test methods*" [70]) are important aspects of the quality of a study.

To assess data quality, the information should be complete and report the necessary type and level of detail. The completeness of the information refers to the conclusion on the comparison between the available information and the information required under the REACH Regulation or the BPR.

A.2.5 Documentation

The adequate documentation of study results is extremely important, both in the study report and in International Uniform Chemical Information Database (IUCLID) robust study summaries. For each endpoint, robust summaries need to be prepared for the key studies (see also ECHA Practical Guide "*How to report robust study summaries*").

Studies conducted according to EU test methods and OECD TGs and under good laboratory practice (GLP) are expected to have adequate documentation (see OECD Principles on GLP). However, when using test methods not following an internationally accepted test method/guideline and/or when the study was not conducted under GLP, special attention should be taken to properly describe the methodology used, the results and interpretation of the results so that the approach can be followed to the extent that the experiment could replicated.

However, in many dossiers in IUCLID, only summaries of studies published in the literature are cited, which does not allow an independent evaluation of the results. For more complex approaches, in particular read-across and grouping, but also for example in the case of the "*defined approaches*" for skin sensitisation published [37], clear guidance on the level of documentation to be provided in IUCLID is available in the ECHA Guidance (for instance regarding defined approaches for skin sensitisation, see Section R.7.5 of the Guidance on IR&CSA – Chapter R.7a).

It is expected that any laboratory with GLP status is able to conduct non-animal approaches under GLP. A GLP quality system helps ensure that the report reflects truthfully the original data. It does not guarantee that the interpretation of the results is scientifically correct. Documentation of the information used under WoE adaptation in a structured way helps to present the results from studies that may have limitations but are considered scientifically correct and ensures that all the information taken together adequately addresses the information requirement with a level of quality and confidence sufficient for regulatory decision making. Similarly, the information used to support read-across must provide confidence that the read-across is indeed acceptable. It is further expected that results from non-validated non-animal approaches are supported with positive and negative control data to increase the confidence in the results and tests should follow good *in vitro* method practice (GIVIMP) for the development and implementation of *in vitro* methods for regulatory use in human safety

assessment [71].

There are publications describing how experiments and their results should be reported (see also ECHA Practical Guides "*How to use alternatives to animal testing to fulfil your information requirements for REACH registration*" and "*How to use and report (Q)SARs*", and OECD "*Guidance document for describing non-guideline in vitro test methods"* [70]). The OECD has also prepared harmonised templates (OHTs) for recording the information used for the hazard assessment of substances. Regarding read-across, the RAAF can be used to structure read-across reporting. There is also a new OECD template for reporting defined approaches [29]. Documentation of use of an individual non-animal approach should not be more complex than that for an animal experiment, however, it requires understanding of the method used.

A.2.6 Mutual acceptance of data

OECD work on harmonisation of test methods (through the adoption of OECD TGs) together with the OECD principles of good laboratory practice (GLP) are integrated parts of the Council Decision on the Mutual Acceptance of Data (MAD). MAD requires OECD countries to accept each other's test data developed for regulatory purposes, if these data were developed in accordance with the OECD Test Guidelines and GLP principles. OECD Member States internationally agree upon test methods for chemical safety assessment and formally adopt such methods as OECD TGs. Before being considered for the OECD TG adoption process, new methods must have been successfully validated in accordance with a standardised procedure (see Section A.2.3 of the present report). There are often different approaches to fulfil a regulatory requirement, which are not all equally accepted in different regions/under different frameworks. MAD provides assurance that a test is not rejected on a quality basis, but does not ensure that a test is accepted as fulfilling a certain requirement.

Currently, there are discussions ongoing on how to promote international acceptance of hazard data resulting from alternative approaches, and in particular from application of IATA. There is as yet no agreement on the inclusion of defined approaches (see Section A.1.3.1) under MAD. However, this may provide a way forward, similarly to what is currently done for OECD TG results. MAD could in principle be **applicable to the predictions** obtained by defined approaches, but it **cannot be applied to the conclusions** obtained by applying IATAs with various undefined methods, and which inevitably includes some WoE determination using expert judgement in the assessment process. OECD has recently included a project in its workplan to investigate the feasibility of TGs for defined approaches under MAD.

A.2.7 Screening vs. fulfilling information requirements

Information on the hazardous properties of substances can be used for different regulatory purposes: basically either for **screening and (de)prioritisation** of the substances to be evaluated or for **fulfilling the information requirements** to support hazard identification/characterisation, hazard classification and/or risk assessment (see also Section A.2.8). Depending on the specific regulatory purpose and whether the information is to be used on its own or together with other data, the level of confidence to be reached may vary and determine the suitability of the information used and, by way of consequence, the suitability of the method used to generate it. Fulfilling the information requirements based on the obligation **to demonstrate the safe use** of a substance requires in general a higher level of confidence in the data than, for example, priority setting for further assessment.

For screening and (de)prioritisation, it is essential to determine the substances that need to be assessed further due to alerts or concerns. The methods to be used should be quick, inexpensive, and be able to screen large numbers of substances and properties. Many non-animal approaches are well suited and used for these purposes as they provide detailed insights into a particular aspect of the toxicological mechanisms that underpin a hazardous property and they are compatible with high content and high throughput technologies. One

challenge is that complex hazardous properties can arise from a variety of different mechanisms, not all of which are fully understood or adequately covered by the currently available non-animal approaches. Therefore, many non-animal approaches investigating a specific mechanism, which may not necessarily be clearly associated to (specific) adversity, are suitable for screening and (de)prioritisation but cannot predict hazardous effects as needed in regulatory decision making. Some examples of such screening methods are methods measuring enzyme activities, receptor binding or gene expression.

Some non-animal approaches may be able to detect substances causing strong effects better than substances with no or mild effects and/or these methods may only partially address the information requirements because (i) substances causing mild effects may not be detected, (ii) predictions of no effects may not always be relied on, or (iii) the methods may not provide information needed for determining classification. These limitations may be linked to the applicability domain of the test system or the fact that the test method only covers a specific mechanism or effects at a cellular level while a broader investigation, potentially addressed by a more elaborated test battery, may be needed to conclude on "*no effect*". Such non-animal approaches can thus be fit for screening and (de)prioritisation purposes but they are less suitable when used to fulfil information requirements, in particular since the prediction of an absence of effects requires a high degree of confidence. For example, the *in vitro* embryotoxicity tests (micromass method, whole embryo culture and embryonic stem cell test) were validated but cannot be used for meeting the information requirements for developmental toxicity due to several limitations, including the fact that they only partially address the *in utero* development and do not take into account maternal metabolism and kinetics.

Under the REACH Regulation, the level of confidence to be reached for each endpoint is reflected by the information requirements of the respective Annex levels. For instance, some *in vitro* tests, like the Ames test, have a good sensitivity (i.e. they give relatively few false negative results) for a specific toxic MoA and negative results from these tests can be used to rule out that specific toxic effect. However, the Ames test only addresses one genotoxic endpoint (i.e. *in vitro* mutagenicity in bacteria). Therefore, even if a negative Ames test is sufficient to fulfil the information requirement at REACH Annex VII, further studies on other genotoxic mechanisms are still needed at REACH Annexes VIII-X levels (and under the BPR) (see also Section B.6). For some other endpoints, such as skin corrosion/irritation and serious eye damage/eye irritation (under the REACH Regulation and the BPR) and skin sensitisation (under the REACH Regulation), the information is required in a stepwise manner, starting with *in vitro* studies, and only if the results from the *in vitro* methods do not allow to conclude or are not applicable will animal testing be used (see Sections B.3 and B.4, respectively).

Finally, there may be some uncertainties associated qualitatively and quantitatively with extrapolating *in chemico*/*in silico*/*in vitro* data for certain endpoints. Due to these uncertainties, such data are often more acceptable for (de)prioritisation or screening (followed by an additional confirmation step) and for use in WoE than for stand-alone fulfilment of information requirements.

A.2.8 Hazard identification/characterisation, hazard classification and risk assessment

As described in Section A.2.7, suitability and acceptability of the use of non-animal data may depend on the specific regulatory purpose.

Hazard identification aims to identify adverse effects caused by the exposure to a substance and separate them from those effects which are not adverse or are secondary to other toxicity.

Hazard characterisation provides information on no observed adverse effect level (NOAEL) and/or a lowest observed adverse effect level (LOAEL) or benchmark dose from animal studies through different administration routes for defining the acceptable exposure level (see Chapter

R.8 of the Guidance on IR&CSA). These values are used not only for risk assessment but also for the classification and (sub)categorisation for the hazards whose criteria rely on dose levels (e.g. STOT RE).

Risk assessment is done under the REACH Regulation and the BPR. This includes the derivation of reference exposure levels for threshold effects, such as a derived no-effect level (DNEL) under the REACH Regulation and an acceptable operator exposure level (AEL) under the BPR, which the exposure should not exceed. Regarding properties that are normally considered as not having any identifiable threshold, such as genotoxicity and genotoxic carcinogenicity, a derived minimal effect level (DMEL) should be determined instead of a DNEL (for further details, see Chapter R.8 of the Guidance on IR&CSA).

Hazard classification is done under the CLP Regulation and is based on criteria for each hazard class and for the different (sub)categories within a class. Currently, the classification criteria for acute toxicity, specific target organ toxicity, carcinogenicity, reproductive toxicity, respiratory sensitisation, skin sensitisation and germ cell mutagenicity are mainly based on evidence obtained from human and/or animal studies. For human health hazards, data from non-animal approaches such as QSAR data only have a supportive role (for more information on what the CLP Regulation is, and how to classify and label substances, see the Introductory Guidance on the CLP Regulation and the ECHA website).

Recital (27) of the CLP Regulation highlights that adequate non-animal approaches for the purposes of classification should be used:

"*The classification and labelling criteria set out in this Regulation should take the utmost account of promoting alternative methods for the assessment of hazards of substances and mixtures and of the obligation to generate information on intrinsic properties by means other than tests on animals within the meaning of Directive 86/609/EEC as laid down in Regulation (EC) No 1907/2006. Future criteria should not become a barrier to this aim and the corresponding obligation under this Regulation, and should under no circumstances lead to the use of animal tests where alternative tests are adequate for the purposes of classification and labelling*."

For classification and labelling purposes, all relevant (human, animal *in vivo*, *in vitro*, *in silico*) evidence of hazardous properties should always be considered together in a WoE determination (CLP regulation, Annex I, point 1.1.1.), to be subsequently compared with the classification criteria. Existing reliable human data should be given high weight and be preferred over other types of data as a basis for classification. However, when reliable data directly comparable to the classification criteria are not available, it may be challenging to determine how much and what type of information is needed for a WoE to adequately support information requirement and hazard classification (and information requirement), especially since the use of the results from non-animal approaches (such as *in vitro* and *in silico* data) to meet the classification criteria (and for sub-categorisation) is not yet as standardised as that of *in vivo* or human data. Defining robust IATAs and approaches with potency information would help standardise the use of results from non-animal approaches for classification and labelling, ensure consistency in the outcome and greatly enhance the use of such data.

CLP criteria not only involve hazard identification but they also often include the notion of degree of hazard. For some hazard classes, such as acute toxicity, skin sensitisation, specific target organ toxicity and fish toxicity, classification and (sub)categorisation are primarily based on the dose or concentration required to cause a certain level of severe effects in experimental studies, which reflects the potency of the substance. For other hazard classes, classification and (sub)categorisation are not dependent on the dose level at which effects occur, but mainly on the severity of the effects, as for example for skin corrosion/irritation and serious eye damage/eye irritation. For germ cell mutagenicity, carcinogenicity and reproductive toxicity, classification and (sub)categorisation are rather linked to the strength of the evidence for

those effects. Any predictive method or approach used for classification and (sub)categorisation should be able to address these aspects. One limitation of many non-animal approaches is that they may not provide quantitative information that could be linked to the adversity and dose-response relationship *in vivo* on which several classification criteria are based. Another challenge is that, for hazard classification, non-animal data are not given the same weight (or strength) as human and *in vivo* data and cannot cover the whole range of categories that are based on the strength of the evidence. To enhance the use of non-animal data for hazard classification, separate or alternative criteria could be set: criteria based on human/animal data (and related to adverse effects) on the one hand and criteria based on non-animal data (and related to the prediction of adversity rather than that of an exact adverse effect) on the other hand. To define a point of departure for predicting adversity, a tipping point – which would be based on "*no return*" or "*likely no longer adaptive*" effects or similar parameters – would need to be determined for each non-animal approach used to meet these non-animal criteria.

Interestingly, work was initiated in 2015 at the Globally Harmonized System of Classification and Labelling of Chemicals (GHS) level to review international efforts to promote non-animal approaches for classification and discuss how to incorporate these in GHS (ST/SG/AC.10/C.4/2015/13), and by way of consequence in the CLP Regulation. The UN GHS sub-committee review the human health hazard criteria in GHS for all hazard classes for potential inclusion/revision of criteria based on non-animal approaches (including read-across, *in vitro*, SAR, etc.). Discussion in a GSH subgroup started on how the results from *in vitro*/non-animal studies should be used in context of hazard classification regarding health endpoints, starting with skin corrosion/irritation. The alignment of data from many animal and non-animal methods with GHS/CLP criteria is currently a challenge.

The CLP Regulation is not concerned with risk assessment. However, if a substance is not considered hazardous (i.e. it does not need to be classified), no evaluation of the exposure or risk characterisation calculations are needed under the REACH Regulation, although the DNEL/PNEC values still need to be derived. Moreover, hazard classification under the CLP Regulation drives risk management through labelling (so that risk can be avoided and safety measures taken) and some classifications (e.g. CMR substances classified in category 1A or 1B) also directly trigger certain risk management measures under legislation other than the chemicals legislation (e.g. legislation on workplace safety, waste legislation, etc.). Under the BPR, classification as CMR category 1A or 1B is one of the exclusion criteria (Article 5 of BPR). These criteria also cover substances that have endocrine-disrupting properties or are considered as PBT or vPvB. Biocide active substances fulfilling the exclusion criteria should normally not be approved, unless a derogation is possible based on BPR Article 5(2). Under the REACH Regulation, the identification of endocrine disruptors and PBT or vPvB substances potentially leads to the inclusion in Annex XIV (list of substances subject to authorisation).

The intended regulatory use may be indicated in the EU test methods and OECD TGs themselves. These standard test methods and guidelines often focus on risk assessment and they may not always address specific aspects of classification and labelling. In addition, EU test methods and OECD TGs aim to be applicable under different regulations and thus remain general. Therefore, specific aspects should be considered to ensure that the results from some of these methods and guidelines will be fully adequate for the intended purposes. ECHA Guidance should be checked for possible specific requirements related to the use of animal tests and non-animal approaches under the REACH, BPR and CLP Regulations. For instance, specific requirements are described in Sections R.7.6 and R.7.3 of the Guidance on IR&CSA – Chapter R.7a for the extended one-generation reproductive toxicity study and for skin sensitisation assessment, respectively.

Some non-animal approaches may be intended for hazard identification only and they may not necessarily be applicable for hazard classification and (sub)categorisation. Information on hazardous properties that is suitable only for classification and labelling and not for risk

assessment may be used for qualitative or quantitative hazard identification only (e.g. skin and eye irritation, skin sensitisation, acute toxicity, germ cell mutation), but without a possibility to derive a NOAEL or a LOAEL (or NOEC/EC_{10} values for environmental endpoints). A qualitative or semi-quantitative risk assessment may still be possible and should be presented. Non-animal approaches used to assess local effects (for example skin corrosion/irritation or skin sensitisation) may be used to derive a safe exposure level for local effects or make a qualitative risk assessment.

For some hazardous properties, like respiratory sensitisation, there are no standard test methods and no standard information requirement under the REACH Regulation or the BPR. The assessment relies on human evidence or a WoE approach using different types of data, such as structural alerts from OECD QSAR Tool Box (v4.0) or lysin-binding properties of the substance as assessed with the DPRA test method (OECD TG 442A). Potential application of non-animal approaches for skin sensitisers to respiratory sensitisers has been discussed [72] and an AOP is under development (AOP 39). Such approaches as mentioned above can be used for hazard characterisation and potentially for (qualitative) risk assessment.

When several non-animal approaches are used to address an information requirement that is based on an animal study, a WoE adaptation (Section 1.2 of both REACH Annex XI and BPR Annex IV) should be used. If this is the case, a WoE determination will be needed for hazard classification as the results from non-animal approaches may not directly match with the classification criteria which are based on animal and/or human data. Both the WoE adaptation under REACH and BPR and WoE determination under CLP may pose a challenge, especially for higher-tier endpoints, for which the information requirements are quite complex and cover many aspects. The general rule for adaptation based on grouping of substances and read-across (Section 1.5 of both REACH Annex XI and BPR Annex IV) requires that, if a grouping concept is applied, substances must be classified and labelled on this basis. Non-animal approaches predicting reliably also non-classification (no hazardous effects) and are addressing elements of information requirements (provide information on the same aspects than information requirement) would enhance their use.

Taken as a whole, the data used to fulfil REACH and BPR information requirements should be adequate for both risk assessment and hazard classification (including sub-categorisation). Using predictive methods when there is an indication of hazard is acceptable from the regulatory point of view if relevant **sub-categorisation for classification** can be achieved. Careful consideration should be given to the adequate use of validated and controlled (information on negative and positive controls) non-animal approaches to avoid or reduce the possibility of over- and under-classification. The intention is that the use of the standard information (required under the REACH Regulation or the BPR or indicated in the CLP criteria) or the use of results from non-animal approaches adequate to adapt it should lead to similar regulatory outcomes with respect to classification (including sub-categorisation) and risk assessment. Currently, non-animal approaches for higher-tier endpoints are not capable of reliable predictions of hazard identification and characterisation as needed for hazard classification, including sub-categorisation, and risk assessment.

A.2.9 Obligation to use non-animal approaches whenever applicable

The use of non-animal approaches for regulatory purposes may require advanced scientific expertise compared to that normally required for an animal study as there is a need to adequately justify their selection, support their validity and applicability, interpret their results, and provide proper documentation. Moreover, non-animal approaches may or may not be cheaper than the animal tests they are replacing but, in some cases, one non-animal test may trigger another, definitive study, which leads to the risk of "multiple" studies and increasing costs. A duty holder may therefore prefer considering the conduct of an animal study. However, the cost and expertise needed to apply some new methods and techniques usually depend on how established a method is and thus may decrease over time. In addition, non-

animal approaches may be more reproducible and/or more predictive or relevant for human or environmental effects. They may also lead to the replacement of tests that use animals or have a potential to cause a high level of suffering. Most importantly, if there is a validated non-animal approach approved for regulatory use, **it must be used to fulfil the information requirement**, unless a scientific justification can be given as to why it cannot be used.

Under the REACH Regulation, ECHA can request in its dossier evaluation decisions that the information requirements are met. The standard information requirements specify the tests and non-animal approaches that can be required. The requirements for the use of non-animal approaches have been set out, for example, in Annex VI and Annexes VII to X of the REACH Regulation, introductory part. In its decisions, ECHA cannot request the use of a non-animal approach to replace a standard information requirement that is specified as an animal study. However, ECHA informs the duty holders on available acceptable non-animal approaches (see Introduction) and requires that, when proposing an animal test, the registrants must provide information on their considerations of non-animal approaches that could potentially replace the animal test and explain why the animal test is still considered necessary.

ECHA has widely advised registrants that they need to consider available non-animal approaches before deciding on the need to conduct new vertebrate tests to meet the information requirements and also requires registrants to show how they have considered non-animal approaches before submitting a testing proposal for a vertebrate test (see ECHA Press release ECHA/PR/15/13). These considerations, which ECHA makes available to third parties through the public consultation on the testing proposals, are taken into account during the decision making on the testing proposal. ECHA has recommended that registrants keep records of their considerations; see Section 2.1 of the "Practical Guide for SME Managers and REACH coordinators" and Section 1.2 of the Practical Guide "*How to use alternatives to testing in animals*", which states: "*As you are obligated to consider alternative methods, you need to keep records of your considerations to support your conclusion as to why it is necessary to generate information using vertebrate animals. You may be requested to submit your consideration of alternative methods.*"

Under the BPR, there are no testing proposals but the applicant is expected to discuss with the evaluating competent authority how to address the information requirements, including considerations on the generation of information by alternative means not involving tests on animals (see recitals 57 and 59 of the BPR).

To address an information requirement for an *in vivo* study, duty holders can produce information using non-animal approaches and then include an adaptation for the requirement of the *in vivo* study, if the information produced is considered adequate to allow the same regulatory outcome (for hazard classification and/or risk assessment). Under substance evaluation, non-animal approaches can be requested by ECHA if considered relevant to address the concern.

A.3 Summary

The regulatory applicability of non-animal approaches depends on many aspects. The detailed requirements for adequate application of non-animal approaches are endpoint-specific, although the principles are the same. Some non-animal approaches (e.g. QSAR predictions) may be used as stand-alone information for some information requirements but this is not yet possible for all information requirements, especially for higher-tier endpoints. This is due to some limitations in terms of applicability domain, coverage of all the relevant endpoint-specific MoAs or validation/standardisation of new technologies.

Integrated approaches combining several methods could be explored to address more complex

information requirements, such as repeated-dose toxicity or reproductive toxicity, and be used under a WoE adaptation. In such integrated approaches, non-animal approaches should be used as supporting information to facilitate interpretation of the effects and strengthen the prediction of the presence or absence of a hazard. Non-animal approaches can also be used to support available but old data, not necessarily compliant with current standard guidelines, or for which some of the information provided by modern investigations may be lacking.

The challenge of developing reliable non-animal approaches is acknowledged. The limitations and challenges differ from those of animal studies and concern different aspects and/or different magnitudes, depending on the method, use and endpoint in question. Non-animal approaches predicting reliably sub-categorisation and non-classification and addressing elements of information requirements are required.

For some endpoints, *in vitro*/*in chemico* investigations or stepwise approaches allow avoiding animal studies in most cases. Many new non-animal approaches and assays have been developed and a growing number of *in vitro* approaches have been accepted for regulatory use or otherwise validated or evaluated for performance. Many of the non-animal approaches can be used as part of grouping and read-across, WoE adaptation, and to screen properties of substances (e.g. for endocrine properties), although screening is not required under the REACH Regulation or the BPR.

Regarding read-across, which is the most often used adaptation under the REACH Regulation, the main challenges for its acceptable use are the lack of robust scientific justification and acceptable supporting data. It is generally considered that, if data for the same endpoint are available from a lower-tier animal study or studies, they can be used to compare the properties of the target and source substances and provide reasonable scientific support for a robust read-across. For example, if the aim is to read-across results from a 90-day repeated dose toxicity study, then information from 28-day studies for all the substances in the category can be used for a robust read-across. However, if that kind of bridging data is lacking, then other supporting information could be produced, for example to show similar MoAs and identical target organ toxicity between the target and source substances. In all cases, bridging data should provide sufficient confidence that the toxicity of all the source and target substances is similar. Care should be taken that these bridging data are produced using similar methods and variables. This requires an extra effort if the data need to be generated but this has the potential to reduce the number of animal experiments.

The main focus areas identified to enhance the use of non-animal approaches are to increase the confidence/clarity for the applicability of the information produced by such approaches, to further develop endpoint-specific standardised and validated non-animal approaches (e.g. IATAs or defined approaches), and to agree the criteria to standardise the use of non-animal data, for example by defining consensus criteria on how to use such data as supporting information for read-across, and by determining the exact elements needed for an endpoint-specific WoE adaptation. Moreover, the use of information from non-animal approaches in rigid/standardised frameworks, such as the current GHS/CLP criteria, may be challenging. For a full replacement of animal studies, changes in current regulatory approaches as such have been suggested, for instance by the creation of new GHS/CLP *in vitro* MoA hazard classes in addition to existing hazard classes, which would allow classification of many substances based on *in silico* and *in vitro* testing results for a limited number of well-known MoAs [61].

Overall, what is needed are standardised and validated methods and approaches that are internationally agreed, and this means validated in a broader understanding. The current non-animal approaches focus on tiered and intelligent testing strategies where tests designed and validated to reflect endpoint-specific KEs are used in a complementary manner. Such testing strategies can indeed address tissue interactions and metabolism to a certain extent, the regulatory use of which will depend on the assays in question. If fit for the purpose, the respective strategies can then be used to support regulatory decision making within their

validated remit and contribute information for complex endpoints. An inventory of available non-animal methods and approaches at different stages of development or already applicable for regulatory purposes would provide an overview of the tools that can be used on their own or in combination to investigate different types of effects. This could help identify potential gaps in terms of coverage of the relevant MoAs for a specific property and/or in terms of standardisation/validation of existing methods and approaches. Such an inventory would in the end facilitate the development of relevant non-animal methods and approaches and enhance their applicability by better targeting the regulatory needs.

Part B: Specific considerations

This part of the report addresses each of the main endpoints for which information is needed under the REACH, CLP and Biocidal Products regulations.

Each section gives an overview of the possible non-animal approaches that can be used to minimise animal testing, including challenges to their use and future perspectives.

Cross-references to relevant ECHA Guidance Documents – which contain more detailed information and recommendations on the use of these non-animal approaches to fulfil the legal obligations under the REACH, CLP and Biocidal Products regulations – are also provided.

A summary of the endpoint-specific information requirements for the REACH Regulation and the BPR, the basis of the CLP criteria, a list of relevant test methods that can be used to fulfil these information requirements (i.e. mostly EU test methods from the EU TMR and OECD TGs), and specific adaptation rules are presented in separate sections of Appendix 3 to this document.

B.1 Toxicokinetics

B.1.1 Description of the information

The term toxicokinetics describes the fate of a substance in the body following exposure. Usual exposure routes for substances are oral, dermal or by inhalation. Toxicokinetics includes four components: absorption, distribution, metabolism and excretion, which are collectively referred to as ADME. Information on toxicokinetics can be obtained from non-animal approaches and from *in vivo* studies and it can enhance the interpretation of the toxicological findings. A more detailed description of toxicokinetics and ADME can be found in the ECHA Guidance (Section 1.3 of the Guidance on BPR Volume III - Parts B+C; Section R.7.12 of the Guidance on IR&CSA, Chapter R.7c), OECD TG 417, as well as in the published paper from Worth *et al*. [10]. It is to be noted that the terms pharmacokinetics and biokinetics are used interchangeably with toxicokinetics. Biokinetics is also being used to describe biokinetic data and models for occupational applications [73] and to describe the *in vitro* biokinetic environment when using cell and tissue cultures [74].

Knowledge of ADME properties of a substance is crucial for evaluating its toxicity. Toxicokinetic information is also important in grouping substances into categories and read-across. One important aspect is being able to predict toxicity for certain groups of substances, for instance based on the formation of similar hazardous metabolites. For ecotoxicology purposes, toxicokinetics provide useful information to assess the bioaccumulation potential of substances which are more likely to bioaccumulate in air-breathing organisms than in fish (see also Section B.12). Such information has already been utilised to identify SVHCs, and the development of this approach for the use of toxicokinetics in bioaccumulation assessment is in progress. The toxicokinetic behaviour of a substance as derived from available data might make further testing unnecessary, for instance in the case that data show the test substance and its metabolites do not reach a specific target organ.

Many *in vitro* cell-mediated assays do not (adequately) address toxicokinetic aspects. However, more complex systems such as organ-mimicking systems or organs-on-a-chip can potentially address some of these aspects and there are number of *in vitro* systems available to determine ADME parameters [1, 75]. The regulatory use of the results from these *in vitro* systems will be highly case-specific as it depends on the specific parameters covered by each system and the extent to which these parameters reflect or can be used to predict the ADME properties of a substance.

The information requirements and the relevant test methods for the REACH Regulation (which does not require any new animal experiments on toxicokinetics) and the BPR are presented in Section 1 of Appendix 3 to this document. There are no criteria for toxicokinetics under the CLP Regulation, toxicokinetic data are used as supporting data but are generally not required.

B.1.2 Approaches for toxicokinetics information

OECD TG 417 provides the test method for the conduct of toxicokinetic studies either as a stand-alone *in vivo* test or in combination with repeated-dose toxicity studies. Measuring toxicokinetic parameters might be possible not only in dedicated assays but also in the course of other guideline studies from sub-acute to chronic repeated-dose toxicity studies [76].

Information on *in vitro* and *in silico* models for toxicokinetics can be found in Section R.7.12 and Appendix R.7.12—2 "*Prediction of toxicokinetics integrating information generated in silico and in vitro*" of the Guidance on IR&CSA – Chapter R.7c, as well as in the published paper from Worth *et al*. [10]. OECD TG 417 also shortly addresses *in vitro* information.

In vitro studies provide data on only specific, limited aspects of toxicokinetics, such as (liver) metabolism or dermal absorption. Their major advantage is that it is possible to carry out in

parallel studies on samples from the species used in toxicity tests and samples from humans, thus facilitating inter-species comparisons (e.g. metabolic profile, metabolic rate constants). PBTK modelling is much used during drug development to simulate *in vivo* toxicokinetics in humans [77]. It can be also used for other types of substances if enough input data are available (e.g. [78, 79, 80]). There are number of approaches published that combine *in vitro* and *in silico* methods (i.e. PBTK modelling) to predict *in vivo* toxicity (e.g. [81, 82, 83, 84, 85, 86, 87]).

Investigation of potential effects of substances through *in vitro* methods (including HTS and HCMs) would require extrapolation to link the *in vitro* effective concentrations to equivalent doses in animals or humans. This technique is called (quantitative) *in vitro* to *in vivo* extrapolation ((Q)IVIVE) and it uses PBTK modelling and information of substance-specific distribution parameters to calculate oral equivalent doses in animals or humans (in mg/kg bw/day). The uncertainty associated with the prediction depends largely on the amount and quality of available data included in the model. A QIVIVE approach under development allows to extrapolate findings from the *in vitro* methodologies to what they mean for the intact organisms (animal or human) [23], but it is not yet known if it fits for regulatory purposes. However, exposure scenarios are needed to place all kinds of hazard information in the context of risk. An SOT/FDA Colloquium, part of the series "*Emerging Toxicological Science: Challenges in Food and Ingredient Safety*", took place in December 2016 (College Park, Maryland, USA) and was dedicated to *in vitro* to *in vivo* extrapolation in safety assessment (see colloquium material).

High-throughput toxicokinetics (HTTK) performs *in vitro* to *in vivo* extrapolation to predict toxicokinetics from rapid *in vitro* measurements and substance structure-based properties. Blood concentrations or target organ concentrations can also be modelled. The method is assumed to be able to predict steady-state plasma concentrations that might be equivalent to *in vitro* bioactive concentrations [88]. The extrapolation is as good as the input information to the models. Relevant information may be used, together with information on human exposure, for risk assessment. Significant change in current classification practice and approach is needed before *in vitro* data with QIVIVE-approach could be used alone for hazard classification.

Human data on toxicokinetics may be available but new studies should not be conducted for the safety assessment of potentially toxic substances as such studies are ethically inappropriate. Biomonitoring (e.g. the analysis of human tissues for direct or indirect evidence of human exposures to substances) can provide unique insight into the relationship between dose and putative toxicity thresholds established in experimental animals. Biomonitoring data may allow development of reverse dosimetry from *in vitro* assays (e.g. [89]). Biomonitoring studies of occupationally exposed retired personnel provide the possibility to define toxicokinetic parameters of elimination without (or at least minimised) effects of continuous exposure, which may impact data validity.

B.1.2.1 Prediction of ADME properties

Information on **oral absorption** should be used, as it greatly improves interpretation of toxicity data. Several physicochemical factors such as ionisation, molecular weight, particle size, log K_{ow}, water solubility, as well as information about the dosing vehicle and oral toxicity can be used in predicting oral absorption (see more details in Section R.7.12 of the Guidance on IR&CSA – Chapter R.7c).

Solubility parameters (e.g. water solubility) may be estimated experimentally or by using *in silico* models: QSAR models, which establish a relationship between physical-chemical activity (properties) and biological activity, or quantitative structure-property relationship (QSPR) models, which are predictors for a given property.

Oral absorption may be estimated by using QSAR/QSPR models that predict permeability

estimates (in %) or through permeability studies where ability to cross a barrier, for example, a lipid membrane or intestinal tissues, is examined:

- *in vitro* permeation studies across lipid membranes (e.g. PAMPA) or across a monolayer of cultured epithelial cells (e.g. Caco-2 cells, MDCK cells);
- *in vitro* permeation studies using excised human or animal intestinal tissues;
- *in vivo* intestinal perfusion experiments on animals or humans.

However, none of the above methods are validated. Caco-2 cells, although the most promising of these methods, has a number of limitations [90].

There are a number of structural alerts, structure-activity relationships (SARs), and other parameters which may be used for predicting absorption and may also be considered in grouping and read-across for toxicological properties. Probably the best-known principle for drug absorption and bioavailability estimation (the amount of the substance which enters into the blood circulation or is available to cause the effects) is the Lipinski rule of five [91]. The simplicity of that scheme prompted development of many variations, which use slightly different (combinations of) descriptors, and calculation methods for them (e.g. [92, 93, 94, 95, 96], Danish QSAR Database and OECD QSAR Toolbox).

Dermal (skin) absorption is an important parameter in the assessment of biocide active substances under the BPR. In dermal absorption, a tiered approach for the estimation of skin absorption has been proposed [97] (see also Section 1.3 of the Guidance on BPR – Volume III (Parts B+C), and Section R.7.12 of the Guidance on IR&CSA – Chapter R.7c). Initially, basic physicochemical information should be taken into account (i.e. molecular mass and lipophilicity), then, a default value of skin absorption is generally used [98]. The principles described in the OECD Guidance on Dermal Absorption [99, 100], as well as the approach and default values described in the EFSA Guidance Document for dermal absorption [101], should be considered. A flow diagram outlining this tiered approach is presented in Appendix R.7.12—4 of the Guidance on IR&CSA – Chapter R.7c. It describes, for example, what to do when there are no dermal absorption data available, which studies to include in the assessment and how to conclude on dermal absorption depending on the data available.

In vitro methods (EU B.45/OECD TG 428) for dermal absorption may adequately reflect the *in vivo* results and they can be used as a replacement provided that limitations are considered. For instance, concentrations remaining in the skin should be generally considered as absorbed and the sum of concentrations in the receptor fluid and in the skin is reflecting the portion of absorbed substance. The lipophilic substances which have a low solubility in the receptor fluid would give especially low absorption value if the concentrations in the skin are not taken into account. More details are presented in Section R.7.12 of the Guidance on IR&CSA – Chapter R.7c, EFSA Guidance on Dermal Absorption [101], SCCS's opinion [102], and in OECD GD 28 [99].

In silico models (e.g. QSAR methods predicting skin permeability coefficient and maximum rate of flux) and mathematical skin permeation models may prove useful as a screening tool or for qualitative comparison of skin permeation potential [103]. On a case-by-case basis, and if scientifically justified, the use of (Q)SARs/QSPRs may be useful, especially within a group of closely related substances.

Rules for skin absorption, such as Flynn's algorithm and Magnuson's rule, have also been developed, which may provide additional information [104, 105]. There are also several other information sources available to predict dermal absorption potential, for example, OECD QSAR Toolbox, Danish QSAR Database and DERMWIN (EPISUITE), which can estimate the dermal permeability coefficient Kp.

For **inhalation absorption**, several physicochemical factors, such as vapour pressure, particle size, log K_{ow}, water solubility and inhalation toxicity data, can be used in predicting inhalation absorption. The multiple-path particle dosimetry (MPPD) model can predict particle deposition and clearance in the lungs.

Parameters related to distribution may be predicted by QSAR models: percentage substance bound to plasma proteins, fraction bound to plasma protein, fraction unbound to plasma, fraction bound and unbound in tissues, apparent volume of distribution and tissue-blood partition coefficient. In addition, principles have been developed for predicting the penetration of a substance through the blood-brain barrier, for example, the Waterbeem's rules [106].

Based on available data, tissue distribution can be mathematically calculated by PBTK modelling using partition coefficients between blood or plasma and the tissue considered. A recent paper provides a state-of-art overview of the available PBTK platforms [75].

Some composite parameters that can be derived from the concentration-time plot of the substance and/or its metabolites can be predicted with some degree of success: these are the area under the curve (AUC) from the Cmax (maximum concentration in the blood or plasma), tmax (time to reach Cmax), and half-life/rate constant of elimination [107].

In vitro tests can inform on **metabolism** by using isolated enzymes, microsomes, cytosol and microsomal fractions, recombinant enzymes, immortalised cell lines, primary hepatocytes in culture or in suspension and liver slices. Most frequently these materials originate from the liver as this is the most relevant organ for metabolism, however, in some cases preparation from other organs are used for investigation of potential organ-specific metabolic pathways. The HepaRG cell line is considered an important *in vitro* tool and can be a good alternative to human hepatocytes [10]. It is also possible to produce metabolically active hepatocytes from human embryonic stem (hES) cells and hepatic stem cells. Recent progress has also been made in generating hepatocyte-like cells from human induced pluripotent stem cells. As a result of JRC's project on *in vitro* human hepatic metabolic clearance methods, a representative method showed a good intra-laboratory reproducibility [108]. Based on the methods from EURL ECVAM, work for the development of an OECD Guidance Document was proposed. Human skin-derived stem cells can be used as a cell source for e.g. *in vitro* hepatotoxicity screening [109, 110].

QSAR prediction models exist for predominant responsible enzyme, percentage of substance metabolised, percentage of substance excreted in the urine, and clearance by the hepatic route. The following freely available tools can provide information on particular mammalian/human metabolism but also distribution (e.g. through brain-blood barrier): OECD QSAR Toolbox, Danish QSAR Database, Meta-print 2D (only metabolism), and SMARTcyp (only metabolism). For metabolism in fish, see Section B.12.

PBTK modelling can provide information on metabolism [75]. Microfluid devices for multi-organ interactions, so-called "body-on-a-chips", provide the possibility to build up physical systems that mimic PBTK models [111]. These multi-organ systems can simulate human metabolism, e.g. biotransformation of a substance to its metabolites [10].

Excretion can be predicted to some extent by relying on physicochemical properties of the molecule. However, depending on the metabolic changes that may occur, the substance or its metabolites that are finally excreted may have few or none of the physicochemical characteristics of the parent substance.

B.1.3 Challenges related to the development and application of non-animal prediction methods

The main hurdle to predict ADME parameters from chemical structures is probably the

complexity of the fate of a substance in the organism as a result of multiple interrelated processes. Availability and accessibility of computer-based tools and reliable databases, as well as availability of well-characterised *in vitro* methods that can generate reliable and relevant ADME data, are a necessary prerequisite for modelling of ADME properties. PBTK models, in general, are considered quite complex and requiring mathematical and programming expertise. Another issue relates to scarcity of substance-specific input parameters for many substances: even if such data are available they often concern drugs or pesticides but not industrial chemicals. For further limitations and challenges, see e.g. [112].

Development and implementation of non-animal approaches able to predict rapidly and at low cost ADME properties has become an important task in assessing the toxicokinetic profile of substances.

B.1.4 Future perspectives

There are a number of issues that should be addressed to facilitate the use of toxicokinetic modelling, such as:

- setting up of databases that collect and store measured parameters and the development of computational models and tools that can make use of these parameters;
- development of *in vitro* tools for high-throughput for measurement of ADME properties for later use in modelling;
- further improvements on the application of analytical methods to measure substances in physiological media.

There are currently several projects ongoing worldwide to address these issues. It is expected that information on toxicokinetics will play a significant role in the future in evaluating species differences and assessing internal exposure [1, 113]. Proper information on toxicokinetics enhances also the use of non-animal approaches for toxicity endpoints.

A QIVIVE approach under development allows to extrapolate findings from the *in vitro* methodologies to what they mean for the intact organisms (animal or human) [23], but it is not yet known if it is fit for regulatory purposes.

B.1.5 Summary and conclusion on toxicokinetics

Information on toxicokinetics is necessary to fully evaluate the toxicological properties of substances. For REACH information requirements, only available information on toxicokinetics needs to be collected but for biocide active substances further information may need to be generated. Toxicokinetics is not a hazard class in the CLP Regulation but it provides useful information for other (toxicological) hazard classes.

A well conducted *in vivo* toxicokinetic study may give important information for subsequent application of non-animal approaches such as application of read-across and building of categories, adapting or triggering further testing. It is recommended to combine investigations on toxicokinetics with other *in vivo* studies that need to be performed to fulfil the regulatory requirements.

B.2 Acute toxicity

B.2.1 Description of the information

Acute toxicity relates to the adverse effects occurring during an observation period of two weeks following oral or dermal administration of a single dose of a substance or a mixture, or multiple doses given within 24 hours, or an inhalation exposure of four hours (see the CLP Regulation and Section 3.1 of the Guidance on the Application of the CLP Criteria). Acute toxicity is characterised in terms of lethality or serious adverse health effects indicative of lethality within a certain observation period. In animal experiments, the adverse effects can be seen as lethality/mortality, or as clinical signs of toxicity, and/or as pathological changes in organs and tissues. Traditionally, acute toxicity of a substance has been characterised by and expressed as a lethal dose where half of the test animals die after administration of a single dose (e.g. an acute oral lethal dose LD_{50}, or an acute inhalation lethal concentration LC_{50} value). However, classification for acute toxicity can also be based on human evidence which shows lethality following exposure.

There are two hazard classes for toxicity resulting from exposure to what is effectively a single dose of a substance: acute toxicity and specific target organ toxicity – single exposure (STOT SE). These classes are independent from each other and both may be assigned to a substance or a mixture if the respective criteria are met (see Sections 3.1 and 3.8 of the Guidance on the Application of the CLP Criteria). STOT SE should be considered where there is clear evidence of toxicity to a specific organ, especially when it is observed in the absence of lethality.

For acute toxicity, classification must be considered for each route of exposure. Substances can be allocated to one of four toxicity categories based on acute toxicity values expressed as (approximate) LD_{50} (for the oral and dermal routes) or LC_{50} (for the inhalation route) values or as acute toxicity estimates (ATE) and according to the CLP criteria.

The information requirements and the relevant test methods for the REACH Regulation, the BRP as well as the basis of the CLP criteria are presented in Section 2 of Appendix 3 to this document.

B.2.2 How to minimise vertebrate animal testing

B.2.2.1 Specific adaptation rules

When met, some of the specific adaptation conditions contained in Column 2 of REACH Annexes VII-VIII or column 3 of BPR Annex II give a possibility not to perform new testing on vertebrate animals to fulfil the information requirement for acute toxicity (for further details, see Section 2 of Appendix 3 to this document).

B.2.2.2 Replacement, reduction and refinement methods

Currently, there are no replacement methods developed for testing acute toxicity. The general rules for adaptation of REACH Annex XI and BPR Annex IV apply, such as read-across and WoE adaptations. A WoE approach, which applies to low toxicity substances and might enable an adaptation of *in vivo* testing of acute toxicity, is presented below.

Grouping and read-across is applicable if the non-animal approach fulfils the criteria described in REACH Annex XI (See also Section R.7.4.4.1 of the Guidance on IR&CSA – Chapter R.7a). QSAR methods may help in grouping and also in supporting the read-across. A few freely available (Q)SAR models and expert systems are capable of predicting acute toxicity and providing relevant and reliable (adequate) data (Section R.7.4.3.1 of the Guidance on IR&CSA – Chapter R.7a).

It could be considered that reduction has been successfully applied for this information requirement since **the current test methods** for the oral route (Acute Toxic Class Method EU B.1 tris/OECD TG 423; Up-and-Down Procedure OECD 425 and Fixed Dose Procedure EU B.1 bis/OECD TG 420) are considered as reduction methods compared to the old one (EU B.1/OECD TG 401) which is no longer in use. Along the same line, OECD TG 402 (EU B.3) for acute dermal toxicity has recently been replaced by a version with a fixed dose procedure, which allows to reduce the number of test animals. There are two OECD-approved acute inhalation toxicity tests, which provide a possibility to reduce the number of test animals, i.e. OECD TG 433 and EU B.52/OECD TG 436 (compared to EU B.2/OECD TG 403). In addition, besides mortality, OECD TG 433 and EU B.52/OECD TG 436 rely on evident toxicity, i.e. "*animals obviously in pain or showing signs of severe and enduring distress*".

When applicable, the use of a **limit dose/concentration** can reduce the number of test animals.

In Section R.7.4 of the Guidance on IR&CSA – Chapter R.7a, ECHA provides advice on how a WoE approach can be used as an adaption of the standard information requirement for acute oral toxicity. The main element of this **WoE adaptation** is a prediction from the sub-acute toxicity study. Based on a IUCLID analysis with almost 10000 registered substances with pre-defined criteria, ECHA has found that low acute oral toxicity (i.e. for substances not to be classified for acute toxicity) can in the majority of cases be reliably predicted from the results of oral sub-acute toxicity studies. For substances with a high NOAEL (at or above 1000 mg/kg bw/day) in the sub-acute toxicity study, the LD_{50} for acute oral toxicity was above 2000 mg/kg in 98 % of the cases. Additional sources of information such as cytotoxicity tests (in particular, the neutral red uptake assay), QSAR and physicochemical information need to be provided. The information must be of sufficient reliability and indicate low acute oral toxicity consistently with the results of the sub-acute toxicity study.

B.2.3 Challenges related to the development and application of non-animal prediction methods

While the WoE approach described above can effectively reduce the number of *in vivo* acute toxicity tests required based on low sub-acute oral toxicity, this approach is not applicable to substances shown to be toxic in a sub-acute oral toxicity study and thus expected to be acutely toxic orally.

Moreover, the *in vitro* methods which have been developed for acute oral toxicity have not shown sufficient sensitivity and specificity and therefore cannot be recommended currently for regulatory use. The neutral red uptake cytotoxicity method is an exception, but its applicability is limited to substances of low toxicity. Acute toxicity *in vivo* through the inhalation route seems currently to be the most difficult to predict and consequently the most difficult to address with non-animal approaches.

Also, specific target organ toxicity (e.g. effects on the central nervous system), poses a challenge since the animal-free approaches do not sufficiently cover all the potential MoAs, and therefore may not always provide a sufficiently reliable quantitative prediction of the acute oral toxicity.

B.2.4 Future perspectives

There are batteries of local (Q)SAR models integrating cytotoxicity information and chemistry-based descriptors which can predict acute toxicity of different substances [114]. These are referred to as hybrid expert systems. Integrating *in vitro* and *in silico* approaches is consistent with other frameworks incorporating mechanistic information for various purposes. One issue is the non-availability of tools to assess toxicity by non-oral routes of exposure [114]. Since

most tools have been developed to evaluate the oral exposure route. Another issue is that the existing (Q)SARs often lack mechanistic basis. The mechanisms behind acute toxicity are far from fully known and hence it is rather difficult to develop QSAR models based on descriptors which have a mechanistic relation to this endpoint.

Efforts have been made to study acute toxicity and assess classification of this parameter with *in vitro* studies and develop non-animal approaches [115, 116]. However, there is no clear vision of a non-animal approach (e.g. combination of methods) that would, in the short term, solve the assessment of acute toxicity.

B.2.5 Summary and conclusion on acute toxicity

Progress has been made to allow the evaluation of the acute oral toxicity of substances which show low sub-acute toxicity in a WoE approach, combining information on acute effects from a sub-acute toxicity study with supporting information from non-animal approaches such as the neutral red uptake cytotoxicity *in vitro* method, QSAR model predictions and other potential sources of information. This possibility for adaptation under the REACH Regulation has been addressed in the update of Section R.7.4 of the Guidance on IR&CSA – Chapter R.7a. The same approach could, in principle, also be applied under the BPR for biocide active substances showing low sub-acute oral toxicity, although this would need to be further investigated.

CLP criteria exist for both acute toxicity and STOT SE and the criteria for both hazard classes refer to information from human and/or animal studies while non-animal approaches are used within a WoE determination. In the future, there may also be combinations of non-animal approaches that may provide predictions adequate for regulatory requirements, in particular hazard classification (especially if GHS and CLP criteria are revised), and that can adequately address acute oral toxicity when the predictions are far above the classification cut-off values.

B.3 Skin corrosion/irritation and serious eye damage/eye irritation

B.3.1 Description of the endpoints

These effects are local effects that occur at the site of contact (skin and eye) irrespective of whether a substance can become systematically available. Substances causing local effects after single exposure can be further distinguished as irritant or corrosive substances, depending on the severity and reversibility/irreversibility of the effects observed.

Corrosive substances are those that may destroy living tissues they come into contact with. Skin corrosion results in irreversible damage to the skin following the application of a test substance up to four hours and which occur by the end of observation at 14 days. Skin irritation is reversible damage of the skin following the application of a test substance for up to four hours. There may be also a concern of dermal effect caused by substances, which cause skin dryness, flaking or cracking upon repeated exposure but which cannot be considered as skin irritants.

Serious eye damage is tissue damage in the eye, or serious physical decay of vision, following application of a test substance to the anterior surface of the eye, which is not fully reversible within 21 days of application. Eye irritation refers to changes in the eye which are fully reversible within 21 days of application.

The information requirements and the relevant test methods for the REACH Regulation and the BPR as well as the basis for the CLP criteria are presented in Section 3 of Appendix 3 to this document. Since *in vitro* methods can in most cases be used to meet the information requirements, there is no urgent need to further develop replacement methods for these endpoints, except for the direct identification of Category 2 eye irritants.

B.3.2 How to minimise vertebrate animal testing

B.3.2.1 Specific adaptation rules

When met, some of the specific adaptation conditions contained in Column 2 of REACH Annexes VII-VIII give a possibility not to perform new testing on vertebrate animals to fulfil the information requirements for skin corrosion/irritation and serious eye damage/eye irritation (for further details, see Section 3 of Appendix 3 to this document). There are no such rules under the BPR for these endpoints.

B.3.2.2 Replacement and reduction methods

The REACH and BPR information requirements already include non-animal approaches as a primary approach. In addition, the general rules for adaptation of REACH Annex XI and BPR Annex IV can also apply, such as grouping and read-across and WoE adaptations.

B.3.3 Challenges related to the development and application of non-animal prediction methods

For serious eye damage/eye irritation, *in vivo* testing may still be needed, due to a lack of *in vitro* methods that can be used for direct identification of CLP Category 2 eye irritants. The recently published OECD Guidance Document on serious eye damage and eye irritation IATA [36] provides examples of how to combine multiple *in vitro* methods to improve the predictivity in identifying CLP Category 2 eye irritants.

Substances causing delayed corrosion effects, i.e. slow skin corrosives, may not always be

correctly detected by *in vitro* methods [117].

B.3.4 Future perspectives

For skin corrosion/irritation, the use of *in vitro* methods will provide results that are suitable for classification and labelling in the majority of cases. An OECD IATA for skin corrosion/irritation was published in 2014 [35].

An OECD Guidance Document on an IATA for serious eye damage/eye irritation [36] was published in July 2017 and provides support in terms of how to use *in vitro* and other data to support the correct identification of all ranges of eye hazard potential. Work is also ongoing to address the detection of reversibility of the effects with *in vitro* methods, for example by using histopathological evaluation [118].

Similarly to how the information requirements have been revised in the REACH Regulation for skin corrosion/irritation and serious eye damage/eye irritation, the use of the available *in vitro*/non-animal methods for these endpoints may be increased under the BPR by a more explicit recognition of these methods in the core dataset for biocide active substances.

B.3.5 Summary and conclusion on skin corrosion/irritation and serious eye damage/eye irritation

Due to the sequential nature of the REACH standard information requirements, and irrespective of the annual tonnage of the substance, new data for skin corrosion/irritation and serious eye damage/eye irritation need to be generated with *in vitro* testing if the available methods are applicable for testing the substance in question. Similarly, for biocide active substances under the BPR, the generation of new data for these endpoints should start with suitable *in vitro* assays. Under the CLP Regulation, there are hazard classes and criteria for serious eye damage/eye irritation and skin corrosion/skin irritation. For serious eye damage/eye irritation, the recently published OECD Guidance Document on an IATA for serious eye damage/eye irritation can now be used to support the correct identification of all ranges of eye hazard potential [36].

If the *in vitro* results are adequate for classification and labelling or risk assessment, no further *in vivo* testing is needed. Registrants therefore need to make sure that the *in vitro* test methods chosen is (are) suitable for the test substance (see also Section R.7.2 of the Guidance on IR&CSA – Chapter R.7a, and ECHA Advice on skin and eye irritation testing to help reduce animal tests).

B.4 Skin and respiratory sensitisation

B.4.1 Description of the endpoints

Skin sensitisation, resulting in allergic contact dermatitis, is an important endpoint to assess for all regulations as this is the most common manifestation of immunotoxicity among humans and has great implications for a person as well as for the community. Therefore, it is important to know whether a substance is a skin sensitiser and how potent it is, to adequately control the exposure to it. The main mechanisms leading to skin sensitisation are relatively well understood. In 2012, the OECD published an AOP which describes the biological mechanisms of skin sensitisation initiated by the covalent binding of substances to skin proteins [119]. The key events (KEs) of this skin sensitisation pathway are:

- KE 1: covalent binding of the electrophilic substance to skin proteins;
- KE 2: release of pro-inflammatory cytokines and induction of cyto-protective pathways in keratinocytes;
- KE 3: activation and maturation of dendritic cells, and their migration to the local lymph nodes;
- KE 4: presentation of the chemical allergen by the dendritic cells to naïve T-cells, which leads to their differentiation and proliferation into allergen-specific memory T-cells.

Even though not considered part of the four KEs leading to the adverse outcome, dermal bioavailability (penetration and, if applicable, metabolism) is a prerequisite for a substance to cause skin sensitisation, i.e. the substance needs to reach the viable epidermis in its reactive form.

REACH information requirements for skin sensitisation have recently been revised to cover for the first three KEs and to make non-animal (i.e. *in vitro*/*in chemico*) methods the primary approach. The information requirements and the relevant test methods for REACH, the BPR as well as the basis of the CLP criteria are presented in Section 4 of Appendix 3 to this document.

Respiratory sensitisation is a term that is used to describe asthma and other related respiratory conditions (e.g. rhinitis, extrinsic allergic alveolitis) due to a hypersensitivity of the airways after exposure to a respiratory sensitiser. There is still uncertainty regarding the exact underlying mechanisms. Based on current knowledge, the induction of respiratory sensitisation can occur through inhalation or dermal exposure to the sensitising substance [120, 121]. An AOP for respiratory sensitisation to low molecular weight substances is currently under development at the OECD (see Project 1.20 of the OECD AOPs development programme work plan and AOPwiki 39).

Respiratory sensitisation is not a standard information requirement under the REACH Regulation and it is part of the additional dataset under the BPR (see Section 4 of Appendix 3 to this document).

However, where relevant, in the case that information on respiratory sensitisation is available, it should be included in the technical dossier and used to support classification and labelling according to the CLP criteria for respiratory sensitisers.

There are no formally recognised and validated animal or *in vitro* tests for respiratory sensitisation. However, there may be data available from human observations indicating respiratory sensitisation in exposed populations or other sufficient evidence, including read-across. In the case that information on respiratory sensitisation is available, it should be included in the technical dossier and used to support classification and labelling where relevant.

B.4.2 How to minimise vertebrate animal testing

Note: Since generation of new animal data for respiratory sensitisation is not a standard information requirement, this section mainly concerns skin sensitisation information requirements.

B.4.2.1 Specific adaptation rules

When met, some of the specific adaptation conditions contained in Column 2 of REACH Annexes VII-VIII or column 3 of BPR Annex II give a possibility not to perform new testing on vertebrate animals to fulfil the information requirements for skin sensitisation (for further details, see Section 4 of Appendix 3 to this document).

B.4.2.2 Replacement and reduction methods

The REACH information requirements for skin sensitisation already include non-animal approaches as a primary approach. In addition, the general rules for adaptation of REACH Annex XI and BPR Annex IV can also apply, such as read-across and WoE adaptations.

B.4.3 Challenges related to the development and application of non-animal prediction methods

The main challenges for the development and application of non-animal approaches for skin sensitisation are that:

- a combination of several *in vitro*/*in chemico* methods with potential other information sources must be used;
- currently the results of the adopted and fully validated *in vitro*/*in chemico* methods available cannot be used directly for sub-categorisation in classification and labelling;
- *in vitro*/*in chemico* methods have their limitations with respect to applicability domain – for example, substances with a low solubility and highly cytotoxic substances can be challenging to properly test *in vitro.* The applicability of these *in vitro* methods may be limited in regard to certain multi-constituent substances and UVCBs (for further details see Section R.7.3 of the Guidance on IR&CSA – Chapter R.7a, ECHA Advice on skin sensitisation testing to help reduce animal tests, and Annex II to OECD GD 256 [37]).

Few approaches for skin sensitisation potency estimation have been described in the scientific literature. However, none of them have been formally approved for regulatory use. Furthermore, some approaches require WoE determination using expert judgement to conclude on potency without clear threshold criteria (e.g. [122, 123, 124]). Some further recommendations on the matter can be found in Section R.7.3 of the Guidance on IR&CSA and the OECD Guidance Document 256 "*Reporting of Defined Approaches and Individual Information Sources to be used within Integrated Approaches to Testing and Assessment (IATA) for Skin Sensitisation*" [37] and its Annexes [125, 126].

Currently, *in vivo* test data may still be needed. However, there are many promising methods and approaches under development or under consideration for regulatory use, including some that could be used for potency estimation (Annex I to OECD GD 256 [37]). Various stakeholders, for example, Cosmetics Europe, are working to overcome this hurdle [127]. More recent developments, for example, the use of TIMES-SS classification trees to predict sensitisation potency [128], may be relevant for mono-constituent substances (but not for UVCBs or multi-constituent substances).

For respiratory sensitisation, there are still major uncertainties regarding the underlying mechanisms. Due to the complexity of these mechanisms, no validated or widely recognised *in vitro* test methods specific to respiratory sensitisation are available yet. Several *in vitro* test

methods have been described in the literature (for instance, see Section R.7.3.9.1 of the Guidance on IR&CSA), but more work is needed to assess their reliability, relevance and performance. Attempts to model respiratory sensitisation have been hampered by the lack of a predictive test protocol for assessing chemical respiratory sensitisation. (Q)SAR models are available but these have largely been based on data for substances reported to cause respiratory hypersensitivity in humans. For instance, the OECD QSAR Toolbox software encodes a profiler (set of rules and structural domains) specific for respiratory sensitisation and offers support to the user in grouping substances which share common structural alerts and possibly predicting the respiratory sensitisation potential through read-across (see Section R.7.3.9.1 of the Guidance on IR&CSA). Presence of activity could be predicted from positive predictions. Absence of effect however cannot be predicted from the lack of alert because the lack of alert might be due to the lack of effect or lack of knowledge.

B.4.4 Future perspectives

It is realistic to estimate that, in the near future, potency estimation can be incorporated into the AOP-based approach methods. This would allow sub-categorisation of more substances for classification purposes without the need for animal studies.

New *in vitro*/*in chemico* test methods for skin sensitisation are being considered or are part of the OECD Work plan for the Test Guidelines Programme 2017 [22]:

(1) LuSens Assay (for KE 2): ESAC peer-review completed and published – test method was included into the OECD TG workplan in April 2017;
(2) SENS-IS (for KE 2 mainly, and maybe KE 3): under validation assessment, peer-review to start in autumn 2016 – this test may be useful for potency estimation, ECVAM contacted for organisation of the Peer review by ESAC;
(3) GARD assay (for KE 3): validation ongoing;
(4) Defined approaches test guideline – feasibility study: development of different combinations of data integration from different key events to produce results similar to or even better than OECD TG 429 (LLNA). The aim is to obtain a full replacement strategy by using non-animal test methods (e.g. *in vitro*, *in silico*) without the need to apply WoE determination using expert judgement. The work will focus as a first step on the Defined Approach case studies presented in the OECD GD 256 Annex I. The project has been approved by the OECD in 2017.

An AOP for respiratory sensitisation to low molecular weight substances is currently under development at the OECD (see Project 1.20 of the OECD AOPs development programme work plan and AOPwiki 39).

B.4.5 Summary and conclusion on skin and respiratory sensitisation

The revised REACH information requirements for skin sensitisation allow the use of an AOP-based IATA in which animal testing may not be necessary. For biocide active substances the use of a similar AOP-based approach may be possible as a WoE adaptation. Whether or not animal testing is needed is dependent both on substance-specific properties, which dictate whether the individual methods within an IATA are applicable to that specific substance, and on test method limitations with respect to risk assessment and classification. The hazard class for skin sensitisation currently refers to human and/or animal data. Further developments are ongoing to improve the predictivity of the individual methods as well as combinations of several methods, for example, regarding hazard classification sub-categorisation needs.
For respiratory sensitisation, the assessment is currently based on existing available information, coming mainly from evidence in humans. More work is needed to better understand the underlying mechanisms and develop standard test methods and prediction models.

B.5 Repeated-dose and chronic toxicity

B.5.1 Description of the endpoint

The term **repeated-dose toxicity** comprises the general toxicological effects (including organ weights and histopathological changes in many organs and tissues, clinical signs, and modifications of clinical chemistry and haematology) occurring through various known or more often unknown MoAs and as a result of repeated daily dosing with, or exposure to, a substance for a part of the expected lifespan (sub-acute or sub-chronic exposure) or for the major part of the lifespan, in the case of chronic exposure. Repeated-dose toxicity studies provide information on possible adverse or hazardous effects, their dose-response relationships, and on their reversibility or irreversibility. Furthermore, these studies may provide information on specific manifestations of toxicity (e.g. neurotoxicity, immunotoxicity, endocrine-mediated effects, reproductive toxicity and carcinogenicity), even though they are not specifically designed to investigate these endpoints. Repeated-dose toxicity studies are important information sources for hazard identification (e.g. hazard classification to specific target organ toxicity (STOT RE) and information on toxicity to reproductive organs) and risk assessment (DNEL derivation). Information from repeated-dose toxicity studies may also indicate a concern and a need for further specific studies, for example for carcinogenicity, reproductive toxicity or (developmental) neurotoxicity.

Separate test methods are available for repeated-dose toxicity studies in rodents and non-rodents, and using oral administration, dermal application or inhalation.

- **Sub-acute 28-day repeated-dose toxicity studies** provide information on the possible toxicological effects arising from exposure to the substance during a relatively limited period of the animal's life span.

- **Sub-chronic 90-day repeated-dose toxicity studies** provide information on the possible general toxicological effects arising from a prolonged period of the animal's life span covering post-weaning maturation and growth well into adulthood, on target organs and on potential accumulation of the substance.

- **Chronic toxicity studies** provide information on the possible toxicological effects arising from repeated exposure covering a significant part of the animal's life span. The duration of chronic toxicity studies should be at least 12 months.

The information requirements and the relevant test methods for REACH Regulation, the BPR as well as the basis of the CLP criteria are presented in Section 5 of Appendix 3 to this document.

B.5.2 How to minimise vertebrate animal testing

B.5.2.1 Specific adaptation rules

When met, some of the specific adaptation conditions contained in Column 2 of REACH Annexes VIII-X or column 3 of BPR Annex II give a possibility not to perform new testing on vertebrate animals to fulfil the information requirements for repeated-dose toxicity (for further details see Section 5 of Appendix 3 to this document).

B.5.2.2 Replacement and reduction methods

The general rules for adaptation of REACH Annex XI and BPR Annex IV apply, such as grouping and read-across and WoE adaptations. There is currently no full replacement method for repeated-dose toxicity testing.

Read-across and category approaches are applicable to predict the repeated-dose toxicity properties. It is recommended to support the read-across data with bridging data such as information on toxicokinetics with the focus on distribution in organs and metabolism. For read-across for a 90-day repeated-dose toxicity study, available information from a 28-day repeated-dose toxicity study or similar may provide the necessary support. Read-across for sub-acute repeated-dose toxicity (28-day) can also be supported by QSAR model predictions, *in vitro* and/or "-*omics*" data (see Section A.1.2.4). The methods listed above may not provide adequate support on their own for sub-chronic toxicity (90-day) or other definitive studies.

There are developments of AOPs based on certain mechanisms/MoAs for organ toxicity (see AOP Wiki). Data from methods investigating such AOPs may potentially support the read-across, but these methods and AOPs are at the moment limited and cannot alone be used to sufficiently reliably predict the absence of toxicity for all organs after repeated dosing *in vivo*. Also information from new approaches such as organs-on-a-chip (see Section A.1.2.3) may be useful. Grouping and read-across may be proposed for hazard identification and characterisation based on information from non-animal approaches without the need to conduct new *in vivo* studies.

Currently, there is no single method to fully replace the animal studies required for repeated-dose toxicity. However, **combination of different endpoints** into a single *in vivo* study gives the possibility to reduce the total number of animals used. It should be ensured that such a combination does not impair the validity and the results obtained for each individual study endpoint. Combination with a repeated-dose toxicity study is covered by several standard test guidelines: combined sub-acute toxicity study with the reproduction/developmental toxicity screening test (OECD TG 422), combined chronic toxicity/carcinogenicity studies (EU B.33/OECD TG 453). Specific considerations for combining a repeated-dose toxicity study with an *in vivo* mammalian erythrocyte micronucleus test are provided in OECD TG 474 and combination with other mutagenicity studies is also possible (see Section R.7.7.6.3 of the Guidance on IR&CSA – Chapter R.7a).

Combining a sub-chronic toxicity study with the extended one-generation reproductive toxicity study could also be considered but some specific aspects should be taken into account (see Section R.7.5.6.3.4 of the Guidance on IR&CSA – Chapter R.7a).

All standard test guidelines for repeated-dose toxicity study give the possibility to perform a **limit test** with the testing of only one dose level of at least 1000 mg/kg bw/day for substances producing no observable toxic effects and if toxicity would not be expected based upon existing data. Furthermore, a proper **design of the testing strategy** and definition of protocols may reduce the number of animals required for testing.

WoE adaptation allows to use methods that alone do not provide sufficient information for repeated-dose toxicity. For repeated-dose toxicity, the main aspects to consider are organ toxicity, length of exposure, dose response, and sensitivity and depth of investigation/predictions done. The adaptation should address all elements of the information requirement in question. Non-animal approaches may be used as elements to support WoE adaptation.

***In silico* and *in vitro* methodologies** may be used in screening and (de)prioritisation and to get insight of the potential toxicological properties of a substance, its mechanisms/MoAs and/or its target cells to be used in combination with other information to support the WoE adaptation.

The generic threshold derived from the threshold of toxicological concern (TTC) methodology might provide a reference value to assess the significance of the human exposure [129]. On a case-by-case basis and at levels of human exposure below the TTC, this approach might support an estimation of a low concern and used in (de)prioritisation. However, its use is

rather restricted due to a number of limitations or drawbacks to the application of TTC for regulatory risk assessment that should be taken into consideration, for example, with respect to applicability domain, route of exposure, classification and labelling (see Appendix R.7—1 to the Guidance on IR&CSA – Chapter R.7c ; Appendix 1-4 to the Guidance on BPR Volume III – Parts B+C; [130]). It should be noted that regulatory applicability of TTC has not been fully investigated for REACH and BPR purposes (e.g. to support exposure-based waiving).

B.5.3 Challenges related to the development and application of non-animal prediction methods

Currently, there are no *in vitro* methods that have been validated for repeated-dose toxicity testing and accepted for regulatory use. The availability of QSAR prediction models for repeated-dose toxicity endpoints is also limited [131]. The main challenge for the development of non-animal approaches for this endpoint is the complexity of the systemic interactions and effects involved in repeated-dose toxicity and the difficulty to model the underlying processes and reproduce integrated responses. Although it may be possible to predict organ toxicity through a certain MoA, it is not yet possible to include all the possible MoAs and interactions between various organs and the organism's systemic and local control systems. This complexity is difficult to predict with computational tools and the *in vitro* systems currently available have mainly been developed to detect certain effects in target organs only. Furthermore, these methods show limitations with respect to coverage of repeated-dose toxicity, kinetics, biotransformation and NOAEL derivation.

An ideal replacement method or combination of methods should address the risk assessment and hazard classification similar to the current information requirement (90-day repeated dose toxicity study). Attempts to reliably reflect the human/animal internal effect levels and external exposure would help to use the data for risk assessment and classification for specific organ toxicity (where the hazard categorisation is based on severity of effects and dose levels). How to address hazard classification for sexual function and fertility is more complex as the categorisation is based on strength of evidence of intrinsic property without consideration of dose levels. Relationship to other systemic toxicity may also influence on categorisation. To conclude on no toxicity, uncertainty is higher than to conclude on hazardous effect if the system and organ selection is not complete and does not reflect the complexity of systemic interactions.

Some commercial statistical QSAR models for human liver, kidney and heart toxicity have however been developed in collaboration with the US FDA to be used as a screening tool, based on information from pharmaceutical clinical trials and scientific articles [132, 133, 134].

B.5.4 Future perspectives

Several EU research projects aiming at the development of new non-animal approaches for repeated-dose toxicity have been initiated, e.g. Predictomics, Pulmo-net, Predict-IV, Seurat-1 and EU-ToxRisk (see https://eurl-ecvam.jrc.ec.europa.eu/validation-regulatory-acceptance/systemic-toxicity/repeated-dose-toxicity). Promising developments – for example, organs-on-a-chip, development of further AOPs and use of "*-omics*", *in vitro*, and *in silico* data – are expected to improve the grouping and read-across possibilities. In addition, information on toxicokinetics (e.g. using *in silico* predictions) would be valuable. New systems combining up to three human organoids are already available and the potential to address repeated-dose toxicity in certain organs *in vitro* seems possible in the future.

B.5.5 Summary and conclusion on repeated-dose toxicity

Repeated-dose toxicity is a complex endpoint involving integrated processes at the molecular, cellular, organ and system levels and for which there is limited knowledge of the underlying mechanistic pathways and their interactions. *In vitro* methods have not yet been validated for

repeated-dose toxicity and the regulatory information requirement cannot currently be predicted by methods such as (Q)SAR models or AOPs and "-*omics*". However, there is a possibility to adapt repeated-dose toxicity studies under the REACH Regulation and the BPR. Grouping and read-across is currently the best approach to avoid new animal tests. For read-across, confidence in similar organ toxicities between the target and source substances is needed and toxicokinetic data can also be used as they provide information for example on organ and tissue concentrations. Non-animal approaches could also be used to support grouping and read-across, either alone for sub-acute repeated-dose toxicity or together with *in vivo* bridging data for 90-day and long-term repeated-dose toxicity studies, or in a WoE adaptation.

Information from repeated-dose toxicity studies is mainly used for classification for specific organ toxicity and reproductive toxicity (i.e. for the STOT RE, sexual function and fertility hazard classes). The CLP criteria for these hazard classes refer to information from human and/or animal studies and depends on nature and severity of effects and effective dose levels.

Relevant scientific developments are ongoing to assess organ toxicity (e.g. organs-on-a-chip). These and other methods predicting specific organ toxicity could be potentially used in current regulatory systems to conclude on hazardous effects on certain organs. To conclude on no toxicity includes more uncertainties and may thus require further supporting data.

B.6 Mutagenicity

B.6.1 Description of the endpoint

The aim of **testing for genetic toxicity** is to assess the potential of a substance to induce genotoxic and mutagenic effects. Genotoxic effects refer to alterations of the structure, information content or segregation of the genetic material (DNA) of a cell. They may result for instance from the impairment of DNA repair and/or protection of microtubule integrity and are potentially reversible.

Mutagenic effects refer to changes of DNA base sequence and/or chromosomal structure in a cell that may be passed on to subsequent generations of that cell. These potentially heritable (if germ cells are affected) alterations, called mutations, may affect a single gene or gene segment, a block of genes, or chromosomes and are irreversible. Genotoxicity is a broader term than mutagenicity and, from the toxicological point of view, genotoxicity can lead to mutagenicity if DNA damages result in permanent alterations that can be transmitted to daughter cells.

Different types of genetic alterations affecting the DNA sequence and genome integrity should be assessed under the REACH Regulation and BPR by covering the following genetic endpoints: gene mutation, chromosome structure aberrations (clastogenicity) and chromosome number abnormalities (aneugenicity). However, no individual mutagenicity test currently covers all of these endpoints and adequate coverage requires the use of multiple tests (in a test battery or a stepwise strategy).

Genotoxicity tests do not provide direct evidence of mutation but they can provide an indication of damage induced in the DNA. For instance, indicator tests detect primary DNA damage (i.e. the first in the chain of events leading to a permanent change), but not the consequences of this genetic damage, which may be cell death, DNA repair to its original state or a mutation after misrepair or no repair. These indicator tests may still be useful to some extent for investigating the MoA of a substance or demonstrating exposure of the target cell or tissue to the substance (or its reactive metabolites) and subsequent impact on DNA. As genetic toxicity can lead to cancer, the usefulness of these indicator tests should also be evaluated in terms of carcinogenicity prediction: in that respect, it is noteworthy that the *in vivo* comet assay, which is an indicator test measuring DNA strand breakage, has shown good performance in the detection of rodent carcinogens that are negative in the conventional bone-marrow micronucleus assay and can be recommended as a follow-up *in vivo* test for this particular type of rodent carcinogens [135].

It should be noted that classification of substances under the CLP Regulation for this hazard class is primarily concerned with germ cell mutagenicity, i.e. substances that may cause mutations in the germ cells of humans that can be transmitted to the progeny. However, the results from mutagenicity or genotoxicity tests *in vitro* and in mammalian somatic and germ cells *in vivo* are also considered in classifying substances within this hazard class (see section 3.5 of the Guidance on the Application of the CLP Criteria).

The information requirements and the relevant test methods for REACH Regulation, the BPR as well as the basis of the CLP criteria are presented in Section 6 of Appendix 3 to this document.

B.6.2 How to minimise vertebrate animal testing

B.6.2.1 Specific adaptation rules

When met, some of the specific adaptation conditions contained in Column 2 of REACH Annexes VIII-X or column 3 of BPR Annex II give a possibility not to perform new testing on vertebrate animals to fulfil the information requirements for mutagenicity (for further details,

see Section 6 of Appendix 3 to this document).

B.6.2.2 Replacement and reduction methods

The basic information requirements for mutagenicity are always *in vitro* data. Results from the *in vitro* test battery may trigger *in vivo* follow-up studies if mutagenicity is observed. There is currently no stand-alone method able to fully replace *in vivo* animal testing for mutagenicity. However, partial replacement or reduction, i.e. application of non-animal approaches as part of a testing strategy or for certain types of substances only, can be achieved in different ways.

Read-across and category approaches are applicable to predict *in vitro* and *in vivo* mutagenicity based on existing information from source substances. However, supporting data are needed, which may be challenging in the case of *in vitro* mutagenicity because there is likely no supporting toxicological data that can be used. For supporting the grouping and read-across of *in vivo* mutagenicity, similar *in vitro* mutagenic properties between the source and target substances can be used. Furthermore, information on distribution and metabolism, for example QSAR and toxicokinetic data, may allow evaluation of the target organs and active metabolites and support the read-across.

If *in vivo* testing is to be performed, the **combination of *in vivo* studies** is strongly encouraged whenever possible and when scientifically justified. This concerns either the combination of different *in vivo* genotoxicity studies into a single study employing a few administrations of test substance (e.g. *in vivo* comet assay and *in vivo* micronucleus test [136, 137, 138] or the integration of *in vivo* genotoxicity studies into a repeated-dose toxicity study [137, 139, 140, 141]. The possibility to combine reproductive toxicity testing with *in vivo* mutagenicity testing could also be considered [142]. It should however be noted that the dose levels used in repeated-dose and reproductive toxicity studies are often not adequate (too low) for genotoxicity testing and additional testing and positive control groups often have to be added in these combined studies. Special precautions need to be taken when conducting a study combining two test methods so that the validity of either test method is not compromised, for instance sufficiently high doses should be used.

It is recommended to include cell samples from both relevant somatic and germ cell tissues (e.g. testes) in a single study if a transgenic rodent (TGR) somatic and germ cell gene mutation assay (EU B.58/OECD TG 488) is foreseen to minimise animal use. Adapted sampling times, in particular an exposure duration long enough to ensure validity of the results for the assessment of germ cell toxicity (see OECD TG 488 for details), and appropriate storage of the germ cell samples are needed. Samples will be analysed if there is a positive result in any of the somatic tissues tested.

The use of a **limit test** is mentioned in most standard *in vivo* test guidelines and, when applicable, it can reduce the number of test animals, like in the *in vivo* mammalian erythrocyte micronucleus (MN) test (EU B.12/OECD TG 474), mammalian bone marrow chromosome aberration (CA) test (EU B.11/OECD TG 475), TGR assay (EU B.58/OECD TG 488) and comet assay (EU B.62/OECD TG 489).

Sometimes the number of test groups may be reduced when conducting an *in vivo* genotoxicity study. For instance, both sexes should not be used if testing in one sex only is possible according to the standard test guidelines and when no sex-specific differences in systemic toxicity or bioavailability have been previously observed with the substance under investigation. Furthermore, if the test is performed in a laboratory with substantial experience and historical data, it should be considered whether a concurrent positive control is really needed (see, for instance, EU B.12/OECD TG 474, EU B.11/OECD TG 475, EU B.58/OECD TG 488). Moreover, it should also be considered whether a concurrent negative control for all time points (*e.g*. for both the 24h and 48h time point in the *in vivo* micronucleus test) will really be necessary ([143]; EU B.12/OECD TG 474).

The information should be evaluated in a separate WoE analysis for each test type and each genotoxic endpoint (i.e. gene mutation or chromosomal aberration). Based on the WoE analysis for each genotoxic endpoint, further testing should be considered in accordance with the regulatory requirements. Additional information, such as predictions from *in silico* techniques, toxicogenomics or mechanistic assays, may support the **WoE adaptation**, thus concluding that information suffice for regulatory conclusion.

Predictions using appropriate ***in silico* techniques** (e.g. substance grouping, read-across or (Q)SAR approaches) can help to confirm results obtained in specific tests, or to develop a better understanding of mutagenicity mechanisms. Several *in silico* tools (QSAR models and expert systems) for genotoxicity are available and have been extensively characterised in the scientific literature [8, 9]. Public domain tools include rulebases in Toxtree [144], profilers in the OECD QSAR Toolbox, VEGA and the Danish (Q)SAR database.

The potential of toxicogenomics, (i.e. gene expression profiling after exposure to a potentially toxic substance to predict its MoA) for *in vitro* genotoxicity testing has been extensively reviewed [145, 146]. Some toxicogenomic studies have shown a correlation between the activation of certain cellular pathways and the mode of genotoxic action [147, 148]. These tests could thus be used to generate supporting mechanistic information to improve genotoxicity assessment. However, further optimisation and standardisation are needed before these tests can be more routinely used.

B.6.3 Challenges related to the development and application of non-animal prediction methods

The basic requirements for mutagenicity are currently covered by a standard battery of *in vitro* tests. Although this *in vitro* test battery can be used to predict genotoxicity, they cannot completely mimic what happens *in vivo* to the substance during the processes of absorption, distribution, metabolism and excretion (i.e. toxicokinetics) or with respect to potential DNA repair. Therefore the standard *in vitro* test battery cannot fully replace animal use and confirmatory *in vivo* genotoxicity or mutagenicity studies are needed when valid *in vitro* mutagenicity test results show mutagenicity.

Another limitation of some of the standard *in vitro* tests is that, although they perform well in terms of identification of genotoxic substances (high sensitivity), they suffer from a high rate of "false" (or misleading) positive results (low specificity). Thus, they identify as positive substances that have not been confirmed as mutagenic (or carcinogenic) *in vivo* [149, 150]. Excessive cytotoxicity, high cell passage number, and increased stress (e.g. due to a compromised p53 response pathway) or DNA repair mechanisms in the cells used for *in vitro* testing have been identified as sources of false positives results *in vitro* [150].

Combination of *in vitro* tests, including the standard *in vitro* test battery, is necessary for adequate coverage of all three mutagenicity endpoints. However, increasing the number of *in vitro* tests in the battery happens at the expense of specificity while only slightly increasing sensitivity [150]. The main consequence of the low specificity of the standard *in vitro* test battery is that follow-up *in vivo* studies will be triggered, which in many cases will be negative.

In vitro genotoxicity studies are performed on cell cultures and are generally considered (by the layman) not to require any use of animals. However, all OECD TG-compliant *in vitro* genotoxicity tests are performed according to two parallel conditions, i.e. in absence and in presence of a metabolic activation system (S9). In other words, *in vitro* genotoxicity tests cannot be correctly performed without the use and sacrifice of several animals.

Only relatively few *in silico tools* are freely available for the prediction of *in vivo* genotoxic potential, and further efforts are needed to develop and evaluate such methods.

B.6.4 Future perspectives

Several initiatives have been undertaken over the past years to reduce the need for animal testing to fulfil the information requirements for mutagenicity. In particular, EURL ECVAM has identified several opportunities for improving the current testing strategy, as reviewed in [151]. Some of the outcomes from these works can already be applied while others are still under development.

As underlined above, one challenge regarding *in vitro* mutagenicity test methods is to improve their specificity while keeping sensitivity high. This can be achieved to some extent by following the updated recommendations from the OECD, as mentioned in the Overview of the set of OECD Genetic Toxicology Test Guidelines and updates performed in 2014-2015 [25]. These recommendations take into account the outcome of recent workshops and expert meetings, including advice on reduced maximum testing concentration, choice of the most suitable test cell lines where possible and preferred cytotoxicity measurement method. Accordingly, some of the OECD TGs for *in vitro* mutagenicity testing have recently been revised and are expected to improve the quality and relevance of the *in vitro* data produced and consequently minimise *in vivo* follow-up studies.

To investigate ways of improving the overall performance (i.e. sensitivity and specificity) of the standard *in vitro* test battery, some analyses were made to determine the optimum number of *in vitro* tests to be performed [152], while still covering the information for the three genetic endpoints required under the REACH Regulation and BPR (see Section 6.1 of Appendix 3). On that basis, some authoritative organisations now recommend the use of a core two-test battery (i.e. a bacterial reverse mutation test and an *in vitro* micronucleus test) for *in vitro* genotoxicity assessment because the *in vitro* micronucleus test can be used to assess both clastogenicity and aneugenicity [141, 153, 154, 155]. It should however be noted that such an approach is not yet implemented in the REACH Regulation and BPR.

Another possibility to better understand the performance of the *in vitro* test battery is to investigate the relevance of Ames test results. To this end, an analysis of different patterns of *in vitro* results was done to identify possible categories of Ames positive results that may be irrelevant or signify low risk of *in vivo* genotoxic or carcinogenic potential [156].

New tests are under development with the aim to (i) either improve the predictive capacity and confirm positive results from the standard *in vitro* test battery prior to, or instead of, *in vivo* studies, or (ii) provide additional mechanistic information as supporting evidence.

More physiologically relevant models have been developed, such as the micronucleus test and the comet assay in 3D human reconstructed skin models for dermal exposure and metabolism [157]. Such models have been proposed as follow-up tests if positive results are obtained from the standard *in vitro* genotoxicity testing battery [158, 159, 160]. The hen's egg test system enables metabolic activation, elimination and excretion of the test substance and has also been proposed to test micronucleus induction as a follow-up test for *in vitro* positives [161].

More recently, assays that simultaneously analyse different biomarkers, including cellular responses to DNA damage, as well as overt cytotoxicity, have been developed to provide mechanistic information on the type of biological damage induced by different types of substances [162].

In addition, the following activities of the OECD Work plan for the Test Guidelines Programme 2017 [22] are ongoing:

- the development of a new test guideline for the Pig-a assay, an *in vivo* gene mutation assay that promotes the 3Rs principle;

- work on a Guidance Document on the adaptation of *in vitro* mammalian cell-based genotoxicity TGs for testing of manufactured nanomaterials;
- miniaturised versions of the bacterial gene mutation test.

B.6.5 Summary and conclusion on mutagenicity

In vitro data are the basic information requirements for the assessment of mutagenicity under the REACH Regulation and BPR. The hazard class for mutagenicity under the CLP Regulation refers to germ cell mutagenicity and the criteria are based on *in vivo* data. Mutagenicity type-specific follow-up *in vivo* testing is triggered if a positive result is obtained when performing the standard *in vitro* test battery.

Improvement of the performance of these *in vitro* test methods and/or consideration of supporting evidence coming from newly developed *in vitro* assays or *in silico* tools are thus warranted to add confidence in the results and minimise confirmatory *in vivo* testing. Reliable grouping and read-across and WoE adaptations are important tools to fill data gaps and reduce the need for testing or for targeting testing needs.

When *in vivo* testing cannot be avoided several reduction approaches exist to limit animal use: current efforts are directed towards reduction of the number of testing groups to the minimum, better integration of the assessment of several genotoxicity endpoints and/or target organs in the same study and combination of genotoxicity and repeated-dose toxicity studies.

B.7 Carcinogenicity

B.7.1 Description of the endpoint

The process of carcinogenesis involves the transition of normal cells into uncontrolled dividing cancer cells through a sequence of stages. This can occur in any organ or tissue but rapidly proliferating cells may be more susceptible.

Carcinogenic substances have conventionally been divided into two categories according to the presumed MoA: genotoxic or non-genotoxic. Genotoxic MoAs involve genetic alterations caused by the substance interacting directly with DNA and changing in the primary sequence of DNA or indirectly following interaction with other cellular processes (e.g. secondary to the induction of oxidative stress). Non-genotoxic MoAs are those effects which do not involve direct alterations in DNA sequence but that may influence gene expression, altered cell-cell communication, or other factors involved in the carcinogenic process, including epigenetic changes. Carcinogenic substances can induce cancer by any route of exposure, but carcinogenic potential and potency may depend on the conditions of exposure (e.g. route, level, frequency and duration of exposure).

The objective of investigating carcinogenicity is to identify potential human carcinogens and their potency. Information should be sufficient to determine the organ specificity of the tumours induced and to establish the dose-response relationship. These data should also enable a comparison against the criteria for carcinogenicity as detailed in the CLP Regulation and subsequent classification of the substances (see section 3.6 of the Guidance on the Application of the CLP Criteria). If the substance is carcinogenic then information on the underlining MoA (threshold or not) and its carcinogenic potency (to define a dose descriptor) is needed. In addition to MoA, relevant elements that need to be addressed include the exposure length, various organs, gender specificity, benignity/malignancy, and time to effects (neoplasia).

A conceptual framework that provides a structured and transparent approach to the WoE assessment of the MoA of carcinogens has been developed [163, 164, 165]. This framework should be followed when the mechanism of action is key to the risk assessment being developed for a carcinogenic substance and can be particularly critical in a determination of whether a substance induces cancer through genotoxic or non-genotoxic mechanisms.

The information requirements and the relevant test methods for the REACH Regulation, the BPR as well as the basis of the CLP criteria are presented in Section 7 of Appendix 3 to this document.

B.7.2 How to minimise vertebrate animal testing

B.7.2.1 Specific adaptation rules

When met, some of the specific adaptation conditions contained in Column 2 of REACH Annex X or column 3 of BPR Annex II give a possibility not to perform new testing on vertebrate animals to fulfil the information requirements for carcinogenicity (for further details, see Section 7 of Appendix 3 to this document). In particular, if the substance is classified as a germ cell mutagen category 1A or 1B, a genotoxic mechanisms for carcinogenicity is presumed and information from a carcinogenicity study is normally not required.

B.7.2.2 Replacement and reduction methods

The general rules for adaptation of REACH Annex XI and BPR Annex IV apply, such as read-across and WoE adaptations.

There are several software tools that can help to build a category with the aim to fill the data gaps with existing information related to genotoxicity and carcinogenicity by **grouping and read-across** [8]. For instance, the OECD QSAR Toolbox contains mechanistically-based profilers and databases of experimental data on genotoxicity and carcinogenicity [166]. This software tool also gives the possibility to form a category based on other similarity criteria, like metabolism. Toxmatch [167] is another software tool that encodes several substance similarity indices to facilitate the grouping of substances into categories and the application of read-across. Other freely available QSAR systems are for example, Oncologic (US EPA), T.E.S.T., VEGA and the Danish QSAR database. The latter contains many different QSAR model predictions from different modelling systems (Case-Ultra, SciQSAR, Leadscope and QSAR model majority predictions) for a high number of *in vitro* and *in vivo* genotoxicity endpoints as well as carcinogenicity bioassay predictions on male and female rats and mice. If applied correctly, the grouping and read-across can be used to fulfil the carcinogenicity information requirement (on the basis of section 1.5 of REACH Annex XI or section 1.5 of BPR Annex IV) or to support a conclusion on carcinogenicity using a WoE approach.

Grouping and read-across may be supported by various non-animal approaches and toxicokinetic data (see Sections A.1.4 and B.1). Information on similar systemic distribution and toxicity provides strong basis for grouping and read-across which may be further supported by information on MoA.

When applicable, the use of a **limit dose** can reduce the number of test animals.

Non-animal approaches can be used as part of a testing strategy or as supporting data for **WoE adaptation**. A general strategy for carcinogenicity assessment under the REACH Regulation is proposed in Section R.7.7.13.3 of the Guidance on IR&CSA – Chapter R.7a and an approach for the BPR in Section 1.9 of the Guidance on BPR Volume III – Parts B+C.

In vitro and *in vivo* mutagenicity can be used in a WoE adaptation for carcinogenicity because mutations and/or chromosomal aberrations are strongly associated with genotoxic carcinogenesis. This correlation is stronger with *in vivo* than with *in vitro* mutagenicity data since *in vivo* assays address the ADME properties of a substance, which could play a critical role in carcinogenicity. Somatic cell genotoxicity can be presumed to be indicative for carcinogenicity and serve as a trigger for investigations for carcinogenicity.

The *in vitro* cell transformation assays (CTAs) have been established to predict carcinogenicity by assessing phenotypic alterations in normal mammalian cells, which mimic malignant transformation *in vivo* [168, 169, 170, 171, 172]. CTAs in general have been proposed to be used as part of a testing strategy and/or in a WoE approach in the testing of substances for carcinogenic potential [172, 173, 174]. OECD Guidance Documents describing the test procedures for conducting the *in vitro* Bhas 42 CTA [174] and the Syrian hamster embryo (SHE) CTA [173] are available. However, the OECD concluded in particular that the predictivity of the CTAs for non-genotoxic carcinogens needed further investigation and considered the approval of draft OECD TGs for these assays premature. The applicability domain of each CTA variant and the limitations described in the above OECD Guidance documents should be taken into account when using these assays.

Several **SAR and QSAR models** and commercial and free expert systems have been published in the literature for predicting genotoxicity and genotoxic carcinogenicity [8, 175]. However, the availability of QSAR models for the prediction of non-genotoxic carcinogenicity is rather limited [144, 176].

If *in vivo* testing is to be undertaken, several options exist to keep the **number of animals** used in carcinogenicity tests to a minimum. Several *in vivo* tests exist that can provide useful data on hazard identification, MoA or carcinogenic potency (for further information, see Section

R.7.7.10.1 of the Guidance on IR&CSA – Chapter R.7a and Section 1.9.2.1.2 (b) of the Guidance on BPR Volume III – Parts B+C). Although the level and type of information obtained will differ from those coming from a standard carcinogenicity study, these tests may support the WoE adaptation, grouping and read-across and a conclusion on the need for a standard carcinogenicity study based on a concern.

Data from **non-conventional carcinogenicity studies**, such as short- and medium-term carcinogenicity assays with genetically engineered (transgenic) animals, may be available [177, 178]. Genetically engineered animals possess mutations in genes that are altered in the multi-step process of carcinogenesis, thereby enhancing animal sensitivity to substance-induced tumours and potentially allowing detection of carcinogens in a shorter period of time and with fewer animals than in a standard carcinogenicity study. However, due to a relative lack of validation and uncertainty as to their value, such assays cannot be used as full replacement methods to the conventional lifetime carcinogenicity studies. They may be used to help determine the need for a standard carcinogenicity study (WoE approach/adaptation). Several evaluations of these types of study have been published (e.g. [179, 180, 181]).

For new studies, the use of the **combined chronic toxicity/carcinogenicity studies** is mandatory under the BPR as it combines separate studies for chronic toxicity and carcinogenicity, and is also recommended under the REACH Regulation when those are needed. As described in the OECD GD 116, "*Conduct and Design of Chronic Toxicity and Carcinogenicity Studies, Supporting Test Guidelines 451, 452 and 453*" [182], useful information on repeated-dose toxicokinetics may be generated as part of a chronic toxicity (OECD TG 452) or carcinogenicity (OECD TG 451) study, or the combined chronic toxicity/carcinogenicity study (OECD TG 453), in addition to data from dedicated toxicokinetic studies such as OECD TG 417.

Under the BPR, a **carcinogenicity study in a second species**, normally the mouse, should be conducted. For the purpose of elucidating the mode of action and human relevance when needed, further investigation of carcinogenicity after obtaining the results of the combined chronic toxicity study should be considered on a case-by-case basis giving priority to the performance of mechanistic studies.

B.7.3 Challenges related to the development and application of non-animal prediction methods

There is currently no stand-alone method able to fully replace animal testing for carcinogenicity assessment. Carcinogenicity assessment should address carcinogenic potency of a substance to all organs and tissues. Thus, it is challenging to develop animal-free prediction methods that could cover all the potential target organs and tissues.

Many mechanisms and KEs of the carcinogenic process are still unknown which makes it difficult to model entirely with an IATA-type approach. In addition, ADME is essential for an understanding of carcinogenic properties of a substance and it is challenging to integrate into non-animal approaches. It is to be noted that ADME is normally implicitly integrated into the carcinogenicity QSAR models which are based on animal *in vivo* data.

Another major challenge is that both genotoxic and non-genotoxic carcinogenicity must be evaluated. Some of the major mechanisms behind non-genotoxic carcinogenicity are known [183, 184, 185], but others are still unknown and knowledge of the corresponding cellular and molecular events is insufficient to allow the development of relevant and comprehensive non-animal approaches. While a number of non-animal approaches exist for assessing the former for screening and (de)prioritisation purposes, there are only limited approaches addressing non-genotoxic mechanisms for carcinogenicity.

Finally, many non-animal approaches are optimised for hazard and MoA identification and not

for the evaluation of potency parameters needed for risk assessment.

B.7.4 Future perspectives

Several non-animal approaches, mostly at the research-level phase, are being developed to try to address carcinogenicity testing, and in particular non-genotoxic carcinogenicity.

Carcinogenesis is associated with multiple changes in gene expression, transcriptional regulation, protein synthesis and other metabolic changes. "-*Omics*"-based high-content methods such as toxicogenomics, proteomics and metabolomics to detect a broad array of molecular changes may be used to identify potential carcinogens based on their molecular signature. These methods can be used to develop new test methodologies *in vitro* [147] or *in vivo* [186, 187, 188, 189, 190, 191], and can also be applied to existing test methods like the CTAs [192, 193, 194] to get some insight into the underlying molecular changes. However, further work is necessary before these methods can be considered for regulatory purposes, including the need for further optimisation, standardisation and formal validation.

Other mechanistic methods, including *in vitro* methods using several cell types, are available to study a number of potential non-genotoxic mechanisms, for example, oxidative stress [195] or inhibition of gap junction intercellular communication [196]. However, these methods are still at the research level and focus on molecular mechanistic understanding. They cannot directly be used for carcinogenicity prediction.

Interestingly, an OECD project is ongoing for the development of an IATA for non-genotoxic carcinogens [197]. This could potentially lead to a standard way of assessing these substances for regulatory purposes and to the development of specific TGs.

B.7.5 Summary and conclusion on carcinogenicity

Carcinogenicity is a complex multi-step process. Chemically-induced cancer may result from a number of different pathways or MoAs and this requires (and allows for) a variety of different approaches to carcinogenicity assessment. While potential genotoxic carcinogens are detected by standard *in vitro* and *in vivo* mutagenicity tests, carcinogens that act by non-genotoxic MoAs are more difficult to identify. This is particularly critical for those potential non-genotoxic carcinogens for which comprehensive modern repeated-dose toxicity data are not available or required (e.g. no 90-day repeated-dose toxicity study or similar) because then triggers for investigations for carcinogenicity cannot be detected. Development of non-animal approaches for carcinogenicity assessment is ongoing but is still at an early stage for non-genotoxic carcinogenicity assessment.

Under the BPR, information on carcinogenicity in two species is required. WoE and grouping and read-across adaptations should be considered before conducting animal experiments. Under the REACH Regulation, information on carcinogenicity is required based on certain criteria (indicated in column 2 of the REACH Annexes). If the criteria are met, grouping and read-across is an important tool to fill data gaps for both genotoxic and non-genotoxic carcinogens and to reduce the need for new *in vivo* carcinogenicity testing.

Regarding the hazard class carcinogenicity under the CLP Regulation, the criteria refer to information from humans and/or animals. One way to enhance the use of non-animal approaches is to use them to support grouping and read-across and WoE adaptations/determination.

B.8 Reproductive toxicity

B.8.1 Description of the endpoint

Reproductive toxicity investigates toxic effects on reproductive and developmental processes by which genetic information is transmitted to the next generation. The ability to reproduce is essential to all species for their survival. Information on sexual function and fertility is necessary for assessing the potential hazard of exposure of men and women of fertile age to a substance and potential effects on next generations. Information on prenatal developmental toxicity to evaluate the potential hazard to the developing foetuses is important if pregnant women may be exposed to the substance. Reproductive toxicity refers in this context to that of mammalian species, such as rodents or rabbits, which are used as laboratory animals in standard reproductive toxicity test methods. While these studies are mainly used to predict effects in humans, they can also provide useful information about reproductive effects in mammalian wildlife species.

Reproduction is a complex process and all the current test methods evaluating this endpoint, including the animal studies, have their limitations. This is because reproductive success relies on integrative action of various organs in both sexes, and simultaneous exposure of several generations at the same time. Interplay between the central nervous system and the reproductive organs (gonads and accessory sex organs) is needed to produce healthy gametes and bring them together for fertilisation. Also following pregnancy, this interplay is important for successful embryo and foetal development, labour, parturition, lactation and nursing of the offspring. Furthermore, potential effects on fertility of the offspring may need to be investigated. This prolongs the duration of the study and also increases the number of animals used due to the necessity to produce subsequent generations.

Investigations of developmental toxicity are focusing on adverse effects that disturb the developing organism, causing functional and/or morphological defects. There are specific test methods available to evaluate the prenatal developmental toxicity, i.e. adverse effects caused by exposure to a substance during *in utero* development. Information on prenatal developmental toxicity may be needed in two species to reduce the uncertainty due to species differences. Further information on prenatal and postnatal developmental toxicity, observable during the postnatal period, may be obtained from reproductive toxicity studies in which the development of the offspring is followed after birth. Information on hazardous effects on postnatal development is needed to assess the potential toxicity for developing children.

A wide variety of mechanisms are assumed to play a role in reproduction in an integrated manner, which poses challenges and makes it practically impossible to address all of them in a non-animal test method. However, some MoAs are known to cause, or to be linked to, specific malformations (falling under prenatal developmental toxicity), and to be linked to reproductive toxicity as reproduction is hormonally controlled. Thus, substances with certain endocrine-disrupting properties may be suspected to cause hazardous effects on reproduction, which need to be addressed. Many of the current test methods for reproductive toxicity also measure parameters sensitive to some endocrine MoAs (see Section B.11 on endocrine-disrupting properties).

Specific areas of developmental toxicity, developmental neurotoxicity and developmental immunotoxicity, are discussed in separate sections.

The information requirements and the relevant test methods for the REACH Regulation, the BPR as well as the basis of the CLP criteria are presented in Section 8 of Appendix 3 to this document.

B.8.2 How to minimise vertebrate animal testing

B.8.2.1 Specific adaptation rules

When met, some of the specific adaptation conditions contained in Column 2 of REACH Annexes VIII-X or column 3 of BPR Annex II give a possibility not to perform new testing on vertebrate animals to fulfil the information requirements for reproductive toxicity. In particular, information on reproductive toxicity may not be needed if the substance is classified in one of the most severe categories for carcinogenicity or mutagenicity and appropriate risk management measures are in place. Similarly, if the substance meets the criteria for classification in the most severe categories for reproductive toxicity and the data is adequate for robust risk assessment, further information on reproductive toxicity may not be needed. Furthermore, there is a possibility to adapt the study if the substance shows no toxicity, there is no absorption and no (or no significant) human exposure. For further details, see Section 8 of Appendix 3 to this document.

B.8.2.2 Replacement and reduction methods

The general rules for adaptation of REACH Annex XI and BPR Annex IV apply, such as grouping and read-across and WoE adaptations.

Currently there is no animal-free/*in vitro* method that can provide information equivalent to the information generated by the reproduction/developmental toxicity screening test (OECD TG 421/422), prenatal developmental toxicity study (EU B.31/OECD TG 414), or extended one-generation reproductive toxicity (EU B.56/OECD TG 443)/two-generation reproductive toxicity study (EU B.35/OECD TG 416).

Grouping and read-across may be used to predict the results for a **reproduction /developmental toxicity screening test** (OECD TG 421/422). However, as the reproduction /developmental toxicity screening test may be the only study available with information on reproductive toxicity, the supporting information for grouping and read-across for this endpoint may be scarce. Relevant supporting data in addition to any information on reproductive toxicity could include bridging data such as information on toxicokinetics with a focus on metabolism and distribution to organs and tissues relevant for reproduction and potentially information on MoA.

To predict the **prenatal developmental toxicity**, good bridging data are needed to support the read-across approach between the targets and sources because even small changes in the structure of a substance may change the developmental toxicity. Information on the reproduction/developmental toxicity screening tests (OECD TG 421/422) and/or MoAs may serve as a good basis for a read-across. Information on toxicokinetics, including exposure of the embryo/foetus and *in vitro* whole embryo culture and other *in vitro* or non-test data, such as QSARs related to embryotoxicity, teratogenicity or prenatal developmental toxicity, may also increase confidence in a read-across approach although may not be sufficient as the only supporting data. If information on prenatal developmental toxicity in a first species is already available for target and source substances, that information may support the read-across for a second species. Tools like Derek Nexus and DART profiling schema in the OECD QSAR Toolbox may help identify group members with the same type of alert for teratogenicity (structural malformations).

For applying grouping and read-across to **extended one-generation reproductive toxicity study** results, information from the reproduction/developmental toxicity screening tests (OECD TG 421/422) and, to some extent, from prenatal developmental toxicity studies, is generally needed for both the source and target substances as bridging data. It should be noted that the study design of an extended one-generation reproductive toxicity study should address the properties of all the target substances. Results from QSAR models may be used as

screening or supporting evidence when assessing the toxicological properties by read-across in a grouping approach. Read-across absence of reproductive toxicity from source substances to target substances requires reliable bridging data.

A combination study (combined repeated-dose toxicity with the screening reproduction/developmental toxicity study), i.e. OECD TG 422, can be considered as a reduction as it reduces the number of animals needed by combining two studies, namely OECD TG 407 (28-day repeated-dose toxicity study) and OECD TG 421 (reproduction/developmental toxicity screening test).

Investigations on prenatal developmental toxicity can be combined with a two-generation reproductive toxicity study or other reproductive toxicity study. However, the benefits may not outweigh the increased complexity of the study, thus, combination studies are not usually recommended. If the expectation is that there is no prenatal developmental toxicity at the limit dose (1 000 mg/kg bw/day), then the addition of a one dose group to evaluate the prenatal developmental toxicity is possible but still an additional control group is needed, which leads to the same animal numbers as in a separate limit dose study.

An extended one-generation reproductive toxicity study and a 90-day study may be combined in specific cases but this needs to be considered carefully since a 90-day study is an important information source for defining the study design for an extended one-generation reproductive toxicity study. Therefore, in practice a combination of these studies could be considered if the study design already includes all cohort expansions based on concern.

When applicable, the use of a **limit dose** can also reduce the number of test animals.

The main reproductive aspects and the confidence level compared to information requirements should be considered when a **WoE adaptation** is applied. For the reproduction / developmental toxicity screening test (OECD TG 421/422), the following **elements** should be addressed in a WoE adaptation: mating and fertility (after a two weeks premating exposure durations), pregnancy, lactation, offspring toxicity including lethality (until PND 13), growth, litter size and sex ratio. Information on histopathology of reproductive organs should be provided. Furthermore, certain endocrine-related parameters such as thyroid hormone measurements, nipple/areolae retention and anogenitial distance should be addressed. The relationship between maternal/paternal toxicity and reproductive toxicity should also be considered, as well as aspects related to the extent of the investigations, sensitivity of the studies, species and routes of administration. In OECD TG 422, various elements on repeated-dose toxicity are also investigated which are not measured in OECD TG 421. The statistical power of the reproduction/developmental toxicity screening test (OECD TG 421/422) is not very high and therefore the results include some uncertainty.

The WoE adaptation should address the potential of a substance to induce prenatal developmental toxicity and its relationship to general toxicity. Thus, the following elements need to be addressed: embryonic/foetal deaths, changes in growth, gross external, skeletal and visceral malformation and variation, sex ratio, and placental toxicity. The extent and sensitivity of the investigations also affect confidence in the evidence. Information on the species differences (developmental effects in a second species) and/or consideration of the relevance of the study results for humans are also needed.

The WoE adaptation for a reproductive toxicity study such as an extended one-generation reproductive toxicity is challenging. The WoE argumentation should include considerations related to sexual function, fertility and offspring (developmental) toxicity observable from the peri- and postnatal period to adulthood. In more detail, all the main elements of reproduction that are normally investigated in an extended one-generation reproductive toxicity study should be addressed in the WoE justification. It is to be noted that extensive investigations on endocrine related parameters are also included in this study and therefore endocrine-

disrupting modes of action (MoAs) should also be considered in the WoE. An element that cannot be fulfilled by a WoE approach, and that is not covered in shorter studies, is the long postnatal period in the F1 generation with measurements of sexual maturation and gonad histopathology in adulthood of the offspring. This information can only be obtained after *in utero* and postnatal exposure and the results can currently not be predicted without an animal test. This information is needed to accept a hypothesis and WoE adaptation based on no effects on reproduction. Furthermore, if criteria for inclusion of extension of Cohort 1B, developmental neurotoxicity and/or developmental immunotoxicity cohorts are met, information on those should be included.

Both the two-generation reproductive toxicity study and the extended one-generation reproductive toxicity study can be considered to provide definitive information on reproductive toxicity. However, there are differences between the two studies such as: information on mating of the F1 animals and production of an F2 generation is included by default in the two generation reproductive toxicity study, whereas more parameters to detect endocrine MoAs and the possibility to include developmental neurotoxicity and developmental immunotoxicity cohorts, if triggered, are included in the extended one-generation reproductive toxicity study. Without additional cohorts, extended one-generation reproductive toxicity study is considered as a reduction method. Furthermore, the extended one-generation reproductive toxicity study includes extensive investigations on clinical chemistry, haematology and histopathology of all organs to evaluate the parental toxicity.

Results from **QSAR models** may be used as a piece of evidence in a WoE adaptation, providing the predictions are within the applicability domain of the QSAR model and the QSAR model meets the OECD QSAR validation principles. QSAR models are usually developed to give binary results for developmental and reproductive toxicity. A positive result from a reliable and relevant QSAR model predicting that the substance has properties related to reproductive toxicity, could provide an indication of hazard potential or trigger further testing. For reproductive toxicity, not all the necessary aspects can be covered by a QSAR prediction. Therefore, a negative result is difficult to accept unless there is other supporting evidence. As developmental/reproductive effects are caused by a multitude of MoAs, most of which are unknown or only partially known, it is in general impossible to predict absence of such effects solely based on non-animal approaches relying on structural information with the same reliability as that of the *in vivo* tests. Furthermore, there is generally a lack of test data on substances that could enhance QSAR model training sets for developmental and reproductive toxicity endpoints. Another limitation of QSAR modelling is that dose-response information, for example the N(L)OAEL required for risk assessment, is not provided. Tools that contain modules related to reproductive and developmental toxicity include Danish QSAR Database, VEGA, T.E.S.T., ADMET predictor, CASE Ultra, Discovery studio Accelrys (former TOPKAT), Leadscope models (including several reproductive toxicity endpoints) and TIMES. Of these, the Danish QSAR DB, VEGA and T.E.S.T. can be used for free.

Current ***in vitro* and *in silico* methods** investigate only partially the embryonic development (e.g. rodent *in vitro* cultures) or part of the potential mechanisms/MoA (AOPs). Thus, information from these prediction methods is not appropriate to support the "no effects" hypothesis but may be used to strengthen read-across or substance categories and maybe WoE and possibly other available evidence providing substantiated reasons for concern (e.g. under substance evaluation). Combination of *in vitro* toxicity data and *in silico* PBTK modelling to predict *in vivo* developmental toxicity for certain substances has been published (e.g. [85, 87]).

Regarding non-animal approaches, they can be used to detect adverse effects and conclude on classification (including categorisation) if supporting information is available. These methods do not provide information on maternal toxicity or allow extrapolation to dose levels *in vivo* (extrapolation at which doses the effects would occur in an *in vivo* test) and therefore it is challenging to consider whether the effects would occur at maternally toxic dose levels or

otherwise extremely high dose levels. Non-animal predictive methods do not provide information for risk assessment (e.g. NOAEL value). Toxicokinetic information is needed to support the dose level considerations. Usually information from predictive methods leading to a concern has been used as a trigger to conduct an animal experiment. Attempts have been made to use reversed toxicokinetics predictions to overcome this (see Section B.1 on toxicokinetics) but there is not much experience on this yet.

B.8.3 Challenges related to the development and application of non-animal prediction methods

Reproduction involves multilevel complex physiological events and signalling networks which are not yet fully known, and also currently too complex to be reliably modelled *in silico*, *in vitro* or by using a combination of various animal-free methods. Even though some individual events could be partially modelled, for example, development of the pre-implantation embryo and uterus implantation, current models do not contain all the control and signalling networks available *in vivo* and therefore they do not reliably model the individual events. In addition, to predict the whole reproductive cycle and development of the whole organism is more complex than combining the (prediction of) individual events. A prediction of no adverse effect on an individual event does not prove there is no adversity for the whole organism. Thus, it is difficult to foresee that any animal-free prediction method could fully replace the animal study in the near future. However, animal-free prediction methods may support grouping and read-across and WoE adaptations.

Although the reproduction process could in theory be divided into several sub-processes which could be modelled, there are still too many unknown aspects and biological events that make the development of AOPs and of a comprehensive testing approach challenging. Therefore, using non-animal approaches under the grouping and read-across and WoE adaptations is currently realistic only as supporting information and cannot be used to conclude on the absence of hazardous/adverse effects, for risk assessment and hazard classification categorisation.

There are developments of AOP-based methods for prenatal developmental toxicity; however, they are MoA-specific and cannot cover all potential developmental hazards. Prenatal developmental toxicity may be caused by various mechanisms and MoAs and by combinations of them which are not yet all known. In addition, the interplay between the developing embryo, the maternal organisms and the placental function involves complex physiological events and signalling networks that are not yet fully known. Therefore, negative results from AOP-based methods cannot provide enough confidence for a "no effect" assumption. Furthermore, hazard classification with categorisation is not yet possible with these methods that are therefore currently not acceptable from a regulatory point of view.

It is also to be noted that the classification criteria of hazard categories for reproductive toxicity refer to evidence from human and animal data. The criteria for Repr. 1B and Repr. 2 state: *"The classification of a substance in Category 1B is largely based on data from animal studies. Such data shall provide clear evidence of an adverse effect"* and *"Substances are classified in Category 2 for reproductive toxicity when there is some evidence from humans or experimental animals".*

To conclude, evidence coming from data without any animal information cannot be used alone for classification purposes.

B.8.4 Future perspectives

It is foreseen that, in the near future read-across and WoE adaptations combining *in vivo* studies and non-animal approaches may be useful options, depending on the substance-specific aspects. Supporting data for read-across may consist of information from a *in vivo*

study such as a reproduction/developmental toxicity screening test (OECD TG 421/422), toxicokinetic data, *in vitro* test results and/or *in silico* predictions. Stem-cell based assays which can mimic some developmental aspects, such as the embryonic stem cell test (and its variations) for cardiac or neural differentiation, may provide a useful model for these effects. Currently, prediction methods that may be used for substances which are strong developmental toxicants, are under development and can also provide useful information for risk assessment. If it is expected that if a group of structurally similar substances cause malformations through a known MoA, it may be possible to use a read-across approach supported by an AOP developed for that MoA. It is possible that new prediction methods for developmental toxicity will be able to provide sufficient information to be used in a WoE approach/adaptation to cover the information in a second species (potential species differences), especially when toxicodynamics and toxicokinetics can be shown to not differ between species.

Some potential adverse effects are not currently investigated by default. These include for example effects occurring later in life which may be caused by early life exposure, long-term (low-level) exposure or potential transgenerational effects. Such effects may be related to hormonal or metabolic changes and may cause premature reproductive senescence, cancer or other diseases. Effects that may occur a long time after exposure are called "latent" and are very difficult and resource demanding to investigate. Many of these effects may be caused by mechanisms and MoAs linked to epigenetics, i.e. permanent or long-term changes in gene regulation affecting the response and cell function. The science around epigenetics is not yet sufficiently developed and investigations cannot yet be included into test methods, but the methodology and interpretation of results are evolving.

Although complex, development of AOPs for reproductive toxicity could help structure the evaluation of the process of reproduction and promote the development of non-animal approaches. Microfluidic platform systems has been developed to simulate organ-organ interaction of hormonal signalling as a phenocopy of menstrual cycle and pregnancy-like endocrine loops. Organ nodules for the ovary, fallopian tube, uterus, cervix and liver have been included mimicking a 28-day human menstrual cycle. The work indicates that tissues of the female reproductive tract, as well as peripheral organs can be integrated into a microphysiologic, dynamic, and microfluidic culture system [198].

B.8.5 Summary and conclusion on reproductive toxicity

There is no acceptable single replacement method available for reproductive toxicity.

The reproduction/developmental toxicity screening test (OECD TG 421/422) is an important information source for substances for which more comprehensive information on reproductive toxicity is not required. This study may provide relevant information for both classification (if hazardous properties are observed) and risk assessment, although it is recognised that the quality of the evidence is less reliable than that obtained through definitive studies. More comprehensive information on reproductive toxicity than the reproduction/developmental toxicity screening test (OECD TG 421/422) is needed under the BPR and at higher tonnage levels for the REACH Regulation (e.g. two-generation reproduction toxicity study or extended one-generation reproduction toxicity study). Both studies provide information for classification and labelling purposes as well as for risk assessment. If significant effects occur, the results may provide a basis for the identification of SVHCs or contribute to this identification together with other information.

Prenatal developmental toxicity studies in two species are important information sources for assessing potential developmental defects under both the REACH Regulation and BPR. The information generated in these studies is relevant for classification and labelling as well as risk assessment.

The CLP criteria for reproductive toxicity, both for sexual function and fertility and for developmental toxicity, refer to information from humans and/or animals and information from non-animal approaches can be used together with other data within a WoE determination. Categorisation depends on the strength of the evidence.

Instead of conducting new animal studies, it may be possible to use existing information alone or supported by production of new (non-animal) information. Reliable information from humans is usually not available as humans are exposed to multiple substances and there are confounding factors, making it difficult to evaluate the effects of single substances.

Regarding reproductive toxicity, to read-across information for extended one-generation reproduction toxicity and/or prenatal developmental toxicity, information on screening studies (OECD TG 421 or 422 or similar) for source and target provide necessary support to the read-across. Limited support can be also obtained for example from *in vitro*, *in silico* and "*-omics*" data, as well as toxicokinetics where this can be provided. Toxicokinetics information showing similar distribution and concentrations in organs, tissues and foetuses/pups may support the grouping and read-across.

Similarly, WoE adaptations may be applicable to avoid new animal studies. These adaptations should address the key elements of the respective endpoint and may combine relevant endpoint-specific information from humans, animals and non-animal approaches. A WoE adaptation for extended one-generation reproductive toxicity study should address all the key elements of the study including the potentially triggered expansions. For such a complex information requirement it may not be possible to build a credible WoE adaptation justification for a no-effect hypothesis. This may be challenging as addressing potential effects in F1 (and F2) generations can currently be done only in an animal study. However, reliable data showing effects meriting classification in Category 1A or 1B for reproductive toxicity could be used to adapt the information requirement.

For substances with expected low toxicity, a limit dose method may be used. Combination of studies may be possible to some extend reducing the animal numbers.

With regard to prenatal developmental toxicity, certain AOPs have been identified and internationally acceptable IATAs and defined approaches may be developed based on these. However, information from AOP-based methods cannot be used to conclude on a lack of developmental toxicity, although it can be used to show that a specific MoA is active or not active for a substance. For sexual function and fertility, a microfluidic platform systems may provide relevant information on potential effects e.g. on menstrual cycle and pregnancy-like endocrine loops. However the method is still under development.

In conclusion, a robust grouping and read-across adaptation with supporting reproduction-related data is currently the most effective way to reduce vertebrate testing. Further scientific development is expected especially regarding prenatal developmental toxicity investigation and detection of endocrine MoAs for reproductive toxicity. One way to enhance the use of non-animal approaches is to use them to support grouping and read-across and WoE adaptations as a separate tests and included in animal studies.

B.9 Neurotoxicity/developmental neurotoxicity

B.9.1 Description of the endpoint

Neurotoxicity studies detect functional changes and/or structural and biochemical changes in the central and peripheral nervous systems. These changes can be of a morphological, physiological (e.g. electroencephalographic changes, biochemical parameters such as neurotransmitter levels), or behavioural by nature.

Currently there is no clear one-to-one link between adult neurotoxicity and developmental neurotoxicity. However, this may be due to insufficient testing as the developing brain is assumed to be more vulnerable than the adult brain depending, on the developmental stage. As an example, one review listed about 200 substances known to be neurotoxic in humans, and just five of these substances have been firmly documented as causing developmental neurotoxicity [199, 200]. Developmental neurotoxicity may cause severe structural and functional abnormalities, including reduction of individual and population intellectual potential (see for example [201]).

The information requirements and the relevant test methods for the REACH Regulation, the BPR as well as the basis of the CLP criteria are presented in Section 9 of Appendix 3 to this document.

B.9.2 How to minimise vertebrate animal testing

B.9.2.1 Specific adaptation rules

Adaptation of information requirements is not relevant because there is no separate standard information requirement for (developmental) neurotoxicity in addition to the general toxicity studies/information, except for certain group of substances where the information is mandatory (such as organophosphates). If there is a particular concern for (developmental) neurotoxicity, information needs may be triggered (REACH Annex IX/X, Column 2), which may lead to the generation of further test data to address the concern. Regarding the BPR, specific information on (developmental) neurotoxicity is part of the additional dataset, and thus, based on concern, similarly to the REACH Regulation.

B.9.2.2 Replacement and reduction methods

The general rules for adaptation of REACH Annex XI and BPR Annex IV apply, such as read-across and WoE adaptations.

Read-across may in general be used as a replacement method when sufficiently reliable and relevant information exist from structurally and (eco)toxicologically similar substances allowing sufficient scientific justification for application of the read-across. Information from other non-animal approaches such as QSAR model predictions and *in vitro* methods may also be used if the source and target substances are within the applicability domain of the model. Currently, however, non-animal approaches are most often used as supporting data to the read-across due to their limitations, and not as stand-alone tests. Because the database of the identified developmental neurotoxicants is limited, the predictability of QSAR models is limited because very few structures associated with developmental neurotoxicity have been identified. For the time being, there is no internationally acceptable specific replacement test method for neurotoxicity or developmental neurotoxicity studies, since, there are no validated non-animal approaches available that are able to predict (developmental) neurotoxicity in humans.

Many existing *in vitro* and *in silico* prediction methods may be used to screen the neurotoxicity potential. Screening methods may provide triggers for further studies, be used as elements in

WoE approaches/adaptation, support read-across and/or provide information on MoAs. Integrated approaches, such as an approach to estimate *in vivo* neurotoxicity integrating *in vitro* neurotoxicity data with biokinetics modelling has been published (e.g. [84]).

Methods using whole organisms provide more complete information because individual KEs are integrated into adverse apical effects and toxicokinetic properties are included. EFSA external scientific review for developmental neurotoxicity [201] presents a potential alternative testing strategy for developmental neurotoxicity based on major KEs tested by a human/zebrafish-based assay.

Methods using whole organisms provide more complete information because individual KEs are integrated into adverse apical effects and toxicokinetic properties are included. A potential alternative testing strategy for developmental neurotoxicity based on major KEs tested by a human/zebrafish-based assay has been presented [201] (see also OECD GD 261 [202] and its Annex 1 [203]). The data from human embryonic stem cell (hESC) differentiation to zebrafish motor behaviour sums up an alternative testing strategy for developmental neurotoxicity covering major KEs of neurodevelopment. Specific, problem-driven research is proposed to be used to fill the current data gaps on neurodevelopmental KEs. The proposed test strategy may be relevant as replacement for developmental neurotoxicity testing but it requires standardisation of individual elements and validation for the total approach. The proposed testing cannot be considered an AOP, although the proposed test strategy uses the term "KEs". It is rather a combination of processes which built on a concept that if the various processes involved in neurodevelopment are not disturbed, then it is likely that the neurodevelopment as a whole is not affected and, in contrast to this, if one process is disturbed, then the neurodevelopment as a whole is affected.

The testing strategy also recommends the use of human-based cell models (preferably derived from induced pluripotent stem cells) as they are considered to have a higher predictability of effects in humans than the rat cell lines. If further standard 28- or 90-day studies are to be conducted, a number of nervous system endpoints are examined and additional parameters could be considered to be included based on concern to potentially reduce further study needs.

To improve identification of substances with potential (developmental) neurotoxicity, integrated testing strategies which combine *in vivo* datasets with *in vitro*, read-across and potential QSAR prediction approaches may be used e.g. within a WoE approach.

The EFSA external scientific review for developmental neurotoxicity lists *in vitro* endpoints (Appendix I in [201]) and individual cell models (Appendix J in [201]). Few *in vivo* model alternatives to OECD TG 426 were identified, but they still need further development and validation. A published evaluation of the developmental neurotoxicity study (OECD TG 426) and corresponding guidance document identifies possible improvements of OECD TG 426 [204].

B.9.3 Challenges related to the development and application of non-animal prediction methods

Test methods predicting developmental neurotoxicity potential of substances faster, less expensive and based on human-specific toxicity pathways would be beneficial for regulatory purposes [205].

For (developmental) neurotoxicity many *in vitro* and *in silico* prediction methods are already available because many MoAs linked to (developmental) neurotoxicity have been identified. However, the justifiable link to adversity and current knowledge about their sensitivity and specificity for (developmental) neurotoxicity, which is a prerequisite for their regulatory use, is still missing.

For instance, cellular models and advanced cellular networks on a chip based on human cells and in 3D configuration simulating organisation of an organ are potentially promising developments. However, these models are artificial systems similar to cell cultures, which do not cover, for example, ADME properties and complex interactions of a living organism [201].

Developmental toxicity testing is challenging due to the existence of critical time windows of sensitivity in the embryonic development (pre- and postnatal) and the complexity of the central and peripheral nervous system. According to Fritsche *et al*. [201], developmental neurotoxicity "*testing of unknown compounds with only 'in vitro' methods may not be sufficient to replace current rodent in vivo testing, especially when (1) the initiating event or primary site of action is unclear or unknown, (2) multiple target sites are hit by chemical compounds, (3) sensitivity for chemicals is dependent of the time window of exposure during early brain development and (4) ADME properties significantly determine biological response".*

There are still knowledge gaps between MIE, cellular and organ effects and the adverse outcome in human developmental neurotoxicity for the most investigated substances such as lead, methyl mercury or PCBs, where there is a strong link between human exposure and adverse outcome. Thus, the chain of events (MoA) is not entirely known even for the most well-known substances with respect to developmental neurotoxicity.

Fritsche *et al*. [201] provided many recommendations, such as: "*Adverse Outcome Pathways (AOPs) for DNT [developmental neurotoxicity] are urgently needed. AOPs help (i) in identification of knowledge gaps. Moreover, (ii) they help in determination if the models (in vivo or in vitro) used for AOP building are – due to their biology – suitable for Key Event evaluation (biological application domain/species differences). Over all, (iii) AOPs will help regulators in using data from alternative approaches in the risk assessment process by creating more certainty"* and "*There is the NEED to obtain experimental data on the clearance index for the placental barrier for a high number of chemicals and pesticides. Once data are available, development of QSAR models for placental permeation can follow, hereby creating a tool for high throughput screening for prioritization of DNT chemicals as part of strategy*".

For some changes it may be challenging to decide whether the effects are adverse or not. However, for regulatory decision making, this is relevant information. A further challenge may be posed by the fact that the classification criteria for specific target organ toxicity after repeated or single exposure and for reproductive toxicity are based on information from humans or studies in experimental animals. Thus, the data from non-animal studies alone are currently not sufficient for classification purposes.

B.9.4 Future perspectives

For gaining regulatory acceptance, definition of biological application domains of non-animal approaches under development by performing for example, *in vitro - in vivo* validation is needed. Moreover, protocols for cell-based and zebrafish assays need international standardisation. With such standardised protocols, the test battery suggested [201, 202] would need to be evaluated for its sensitivity and specificity by testing concentration-responses of known substances positive and negative for developmental neurotoxicity across the different assays.

A need to improve interpretation of information from rodents is high. Knowledge on pharmaco-/toxicodynamics of the developing brain of rodents compared to humans would enhance the interpretation of the results from developmental neurotoxicity studies in human situation. Furthermore, in OECD TG 426 and the corresponding guidance documents, in terms of their reliability and usability for scientific and regulatory judgements and decision, are areas for improvement [204].

Developmental neurotoxicity is an area where there is an urgent need for the development of

non-animal methods and approaches, especially due to the limitations of the information obtained from existing animal studies. Presentations held during SOT FDA Colloquia (2016) and OECD/EFSA workshop (2016) on development of methods for developmental neurotoxicity can be found online. The development and use of new testing methods (*in vitro*, non-mammalian models such as zebrafish, QSAR, etc.) will facilitate the faster identification of developmental neurotoxic substances and could therefore enable the generation of robust databases for developmental neurotoxicity based on which development of read-across tools specific for developmental neurotoxicity will be done (OECD QSAR Toolbox).

Currently, only a few AOPs for developmental neurotoxicity are available, and because the development of AOPs is challenging, the development of a sufficient number of specific AOPs for developmental neurotoxicity will take time. Therefore, not to delay development and implementation of the testing strategy for developmental neurotoxicity, it was suggested during the OECD/EFSA workshop (October 2016) [202, 203, 206] that neurodevelopmental processes can also be utilised as anchors for *in vitro* assays development. Various neurodevelopmental processes are evaluated with a test, or a pattern of tests, and although the full network, feedback and control mechanisms cannot be included in the approach, there is an assumption that if various processes are not disturbed then also the whole neurodevelopmental function is not disturbed either. This concept presented is considered suitable for screening and (de)prioritisation of substances. The implementation of a testing strategy for developmental neurotoxicity should be carefully considered once validated and accepted. At OECD level (project 4.124, [46]) work is started (2017) to develop a new guidance document on developmental neurotoxicity *in vitro* assays.

B.9.5 Summary and conclusion on (developmental) neurotoxicity

Both the REACH Regulation and the BPR require information on (developmental) neurotoxicity if there is a particular concern due to earlier findings, MoAs or structural similarities to other substances known to cause (developmental) neurotoxicity.

There is no specific hazard class developmental neurotoxicity, but neurotoxicity is considered under STOT RE and developmental toxicity under toxicity to reproduction. If there is a concern, a read-across and a WoE adaptation is the most recommended approach to avoid new animal testing. Currently there are no internationally accepted *in vitro* or *in silico* methods, nor information from invertebrate or non-mammalian animals, which can be applied alone to replace an animal study in mammals. However, needs and developments in the area are recognised and many methods can serve to enhance a read-across approach, as elements in a WoE approach/adaptation or in screening and (de)prioritisation of substances. There is a need to further standardise and validate the various approaches, including the potential alternative developmental neurotoxicity testing strategy as described in [201], OECD GD 261 [202] and its Annex 1 [203].

Once standardised and validated, the approaches under development could be used to screen and (de)prioritise the substances for potential further assessment or included in testing strategies by the registrants as a first step for testing under substance evaluation. One way to enhance the use of non-animal approaches is to use them for screening and (de)prioritisation. Currently, the performance standards and readiness criteria for individual *in vitro* assays for developmental neurotoxicity are under development. This effort will lead to development of an OECD GD on available *in vitro* test methods for developmental neurotoxicity used alone or in combination (e.g. within an IATA) for various regulatory purposes. Development of a Guidance Document on *in vitro* assays for developmental neurotoxicity has been included in the OECD Work Programme 2017 [22] and will be established in collaboration with EFSA, European and USA experts.

B.10 Immunotoxicity/developmental immunotoxicity

B.10.1 Description of the endpoint

Immunotoxicity investigates changes in innate and adaptive immune responses caused by substances. Immune dysfunction may have severe health consequences such as immunosuppression and reduced resistance to infections, or exaggerated immune responses such as allergy and autoimmunity, or inflammatory-based diseases or pathologies, for example, tumours.

Immunotoxicity is a potential systemic toxicity considered in a general risk assessment [207]. Assessment of immunotoxicological risk relies on a variety of endpoints that reflect immune system health and the differences in the approaches to assess the risk of immunotoxicity. In addition, other forms of toxicity are minimal but stem from a lack of basic knowledge and scientific understanding of the immune system endpoints that link cellular toxicity with downstream disease outcomes [207].

Similarly to the reproductive and central nervous systems, the immune system is particularly vulnerable to substance exposure during development, and function declines with age, resulting in increased risk of adverse health outcomes from substance exposure at the extremes of age [207]. Generally, functional investigations are considered more sensitive than non-functional measurements.

The information requirements and the relevant test methods for the REACH Regulation, the BPR as well as the basis of the CLP criteria are presented in Section 10 of Appendix 3 to this document.

B.10.2 How to minimise vertebrate animal testing

B.10.2.1 Specific adaptation rules

As there is no separate standard information requirement for (developmental) immunotoxicity in addition to the animal studies with the focus on other endpoints, adapting the test is not relevant. If there is a particular concern for (developmental) immunotoxicity, information needs may be triggered, which may lead to the generation of further test data to address the concern (REACH Annex IX/X, column 2). Regarding the BPR, specific information on (developmental) immunotoxicity is part of the additional dataset and thus based on concern, similarly to the REACH Regulation.

B.10.2.2 Replacement and reduction methods

The general rules for adaptation of (REACH Annex XI and) BPR Annex IV, such as read-across and WoE adaptations.

Read-across may in general be used as a replacement method when sufficiently reliable information exists from structurally and (eco)toxicologically similar substances allowing sufficient scientific justification for application of the read-across. Information from other non-animal approaches such as QSAR model predictions may also be used if the predictions are within the applicability domain of the model. Currently, however, non-animal approaches are most often used to support the read-across and not as stand-alone tests due to their scientific limitations.

Currently there is no internationally accepted specific replacement test method for immunotocixity or developmental immunotoxicity studies. Information from immunosuppression tests included in other international guidance documents (e.g. EMA ICH

Topic 8 [208], EPA (OPPTS 870.7800) may also be considered (see also Section 10 of Appendix 3). Due to the lack of EU/OECD methods, enhanced examinations included in repeated-dose toxicity studies may allow a better identification and distinction of effects associated with the immune system (e.g. immunosuppression and -stimulation).

So far there are no validated non-animal approaches available that are able to predict (developmental) immunotoxicity in human. Many existing *in vitro* and *in silico* prediction methods may be used to screen the immunotoxicity potential. Screening methods may provide triggers for further studies, be used as elements in WoE approaches, support read-across, or provide information on MoAs. Methods using whole organisms provide more complete information because individual KEs are integrated into adverse apical effects and toxicokinetic properties are included. If further standard studies (e.g. a 90-day study) are to be conducted, some information on the immune system is gained and additional immune parameters (e.g. functional test aspects) could be considered to potentially reduce further study needs. However, some investigations require inclusion of positive control animals.

Read-across is a powerful reduction method when testing is planned for structurally and (eco)toxicologically similar substances, and should be used whenever possible and reliable. Most often, information from animal studies is used for read-across, but also information from an *in vitro* test may be considered if the target and source substances are within its applicability domain.

To improve identification of substances with potential (developmental) immunotoxicity, integrated testing strategies which combine *in vivo* datasets with *in vitro* approaches may be applied, e.g. within a WoE approach.

B.10.3 Challenges related to the development and application of non-animal prediction methods

For (developmental) immunotoxicity, *in vitro* and *in silico* prediction methods are already available. However, the justifiable link to adversity and their specificity for (developmental) immunotoxicity is still missing and this is needed for regulatory purposes. As with many *in vitro* models, these models often lack interactions with other cells/organs (e.g. surrounding organism) and do not cover ADME properties.

Developmental toxicity testing is challenging due to the existence of critical time windows of sensitivity in the development and the complexity of the immunosystem network. The normal maturation of the immune system depends on specific processes that differ in both time and location within the body, and the immune system of the non-adult is a moving toxicological target for xenobiotic interactions [209, 210, 211, 212]. The prenatal, neonatal, juvenile and adolescent immune systems should be viewed as distinct from that of the adult in terms of risk assessment.

A further challenge maybe the classification criteria for specific target organ toxicity after repeated or single exposure, as well as the classification for reproductive toxicity, which is based on information from humans or studies in experimental animals. Thus, currently the data from non-animal studies alone is not sufficient for classification purposes.

B.10.4 Future perspectives

Interpretation and investigations for immunotoxicity and immunomodulation, including non-animal approaches are still under development (see e.g. [213]). The most important aspect is to define the link between the adversity and immunomodulation. Testing potential immunomodulation can be useful only if the results have regulatory impact, i.e. they can be used for risk assessment or classification and labelling purposes. Thus, the threshold for adversity must be set for regulatory uses, which needs further scientific research to support

policy decision.

In case of concern, further testing to investigate immune function should be conducted only if the outcomes of such studies can be interpreted in relation to the risk assessment and/or classification and labelling for the substance of interest.

B.10.5 Summary and conclusion on (developmental) immunotoxicity

Both the REACH Regulation and BPR require information on (developmental) immunotoxicity based on particular concerns due to earlier information on findings, MoAs or structural similarities to other substances known to cause (developmental) immunotoxicity.

There is no specific hazard class developmental immunotoxicity, but immunotoxicity is considered under STOT RE and developmental toxicity under toxicity to reproduction. If there is a concern, read-across and WoE adaptations are the most recommended approaches to avoid new animal testing. Currently there are no internationally accepted *in vitro* or *in silico* methods, nor information from alternative organisms which can be applied alone to replace an animal study. However, needs and developments in the area are recognised and many methods can serve to enhance a read-across approach, as elements in a WoE approach/adaptation or in screening and (de)prioritisation of substances. There is a need to further standardise and validate the various approaches, including the potential alternative developmental immunotoxicity tests.

The current *in vitro* methods and approaches are recommended to be used to screen and (de)prioritise the substances for potential further assessment.

B.11 Endocrine properties/MoAs

B.11.1 Description of the MoA

WHO/IPCS defines **an endocrine disruptor** as "an exogenous substance or mixture that alters functions of the endocrine system and consequently causes adverse health effects in an intact organism, or its progeny, or (sub-)populations." [214].

Information on the endocrine properties of a substance has become important due to the increasing concern for potential effects on health and environment through endocrine MoAs. "Endocrine disruption" is not an apical endpoint in itself but endocrine MoAs may lead to adverse effects such as cancer or reduced reproductive health at individual and/or population levels. Generally, the adverse effects resulting from endocrine disruption are linked to malfunction of an endocrine organ, or impaired transfer of the products of an endocrine organ, including hormones, or interference with the messages received by an endocrine organ (for further explanation for WHO/IPCS definition and OECD Conceptual Framework, see Section 11 of Appendix 3). Many endocrine MoAs are known, but, so far, there are standardised methods to assess only (anti)oestrogenicity, (anti)androgenicity, steroidogenesis and thyroid toxicity [215]. For the purpose of this document, endocrine-disrupting properties of a substance refer to its effects, adverse or non-adverse, that are likely to be due to an endocrine MoA. A substance showing endocrine-disrupting properties is not necessarily an endocrine disruptor in the regulatory sense. The identification of a substance as an endocrine disruptor requires evidence of adverse effects likely due to its endocrine MoAs.

The mammalian toxicity studies that include endocrine MoA parameters are conducted to investigate information requirements such as repeated-dose toxicity, reproductive toxicity and carcinogenicity. However, for ecotoxicity assessment, such studies are only requested to clarify whether a substance is identifiable as an endocrine disruptor.

Further information on current activities in relation to endocrine disruptors at EU level can be found on the European Commission website. UNEP and WHO have published a document, "State of the science of endocrine disrupting chemicals" [216].

The information requirements and the relevant test methods for the REACH Regulation, the BPR as well as the basis of the CLP criteria are presented in Section 11 of Appendix 3 to this document.

B.11.2 How to minimise vertebrate animal testing

B.11.2.1 Specific adaptation rules

As there is no separate REACH standard information requirement for endocrine-disrupting properties or endocrine-disrupting MoAs, adapting specific endocrine-related animal testing is not relevant. However, information may be requested under the substance evaluation process based on concern. Regarding the BPR, information requirements on endocrine-disrupting properties are part of the additional dataset, and thus based on concern, but there is no corresponding specific adaptation rules.

B.11.2.2 Replacement, reduction and refinement methods

The general rules for adaptation of REACH Annex XI and BPR Annex IV apply for studies which include investigations for endocrine properties. Thus, approaches such as read-across and WoE adaptations may use information from non-animal approaches to address those investigations of a (standard) information requirement.

3R perspectives and non-animal approaches might be of relevance for assessing endocrine-disrupting MoAs. However, for the identification as an endocrine disruptor, the assessment of apical effects at organism-level or population relevant effects is needed and requires animal testing. It should be noted that a number of tests in amphibians exist and may provide information on endocrine MoAs such as disturbance in thyroid hormone regulation (see Section 11.3 of Appendix 3 to this document).

For investigating an endocrine MoA, *in vitro* tests and *in vivo* screening tests might also be considered. *In silico* models could provide a rapid screen for prediction of endocrine active substances. To reduce the complexity of building a plausible link from MoA to adverse effects (which is a prerequisite for identification of a substance as endocrine disruptor), validated endocrine-mediated AOPs might help focus the tests on specific KEs and reduce the need for animal testing.

B.11.3 Challenges related to the development and application of non-animal prediction methods

The main challenge for endocrine testing is to have testing systems sufficiently complex to cover all relevant parts of the signalling network. This is a particularly true with respect to hormonal phenotypic plasticity [217].

Because there are many mechanisms and MoAs involved in endocrine disruption, these need to be clearly identified before appropriate prediction methods can be developed. The relevance of these mechanisms and MoAs to human health and wildlife, including their role in the sensitivity of various life stages, needs further investigation to eventually enable the characterisation of the hazardous properties related to each mode of endocrine action. In addition, a certain mode of endocrine-related activity may not necessarily lead to adverse effects, even if such an activity triggers concern for potential adverse effects. Prediction of the occurrence or absence of one endocrine MoA for a substance, even if including measures of adversity, may not rule out other endocrine MoAs and related adverse effects. Therefore, evidence of adverse effects in an intact organism is needed for considering a substance as an endocrine disruptor.

Another aspect to consider is that it is very difficult, if not impossible, to predict long-term effects of early exposure to a substance during development based on currently available *in vivo* and non-animal approaches. Long-term follow-up of the consequences of exposure at any phase of development is also not feasible in mammalian studies. Human data may be available only after a significant amount of health problems have been detected, and the exposure to multiple substances and stressors may not allow to conclude retrospectively on a causal link to any of these. However, it is important to try to address these possible effects that occur only much later in life. To this end, innovative non-animal approaches to predict potential adverse health outcomes and effects on populations need to be developed, together with methods capable of estimating the uncertainty involved. Finally, the minimum level of confidence needed for taking a regulatory decision based on results from these methods should also be determined.

One critical issue is that the current approach relies on the identification of adverse apical effects in intact organisms. Further work on AOPs is needed to help establish a body of evidence that may in the future make a full *in vivo* test not necessary.

B.11.4 Future perspectives

Further scientific and methodological developments in the field are expected, including new AOPs and AOP networks, more knowledge on the endocrine network complexity, new *in vitro, in silico* and other prediction methods, and the inclusion of further parameters in animal experiments to investigate endocrine-disrupting properties of substances and their relationship with adverse effects. There is no particular endocrine disruption hazard class under the CLP

Regulation; the substances with hazardous outcome are classified according to the hazard (e.g. for reproductive toxicity or carcinogenicity). Substances with hazardous properties likely caused by endocrine MoAs may be considered substances of very high concern (SVHCs) under the REACH Regulation.

The criteria for the identification of endocrine disruptors for plant protection products and biocidal products, as well as the corresponding Guidance document, are currently being developed by the European Commission (for further information, see the European Commission Policy, Communication from the Commission to the European Parliament and the Council [218], draft criteria for biocidal products [219], draft criteria for plant protection products [220], and impact assessment [221]).

The question as to how to address potential low-dose effects and cumulative/combined effects based on endocrine MoAs needs further attention. The various mechanisms behind the endocrine MoAs at various life stages and diseases, species differences, and their relevance to humans needs further investigations to clarify the link between exposures, MoAs and diseases and how the findings in various species predict the human health and environmental hazards.

Approaches based on "-*omics*" may prove useful for predicting the endocrine activity of new substances that have not yet been tested in standard animal tests.

B.11.5 Summary and conclusion on endocrine disruption

There are neither specific information requirements nor a specific hazard class for endocrine disruption, endocrine-disrupting properties or endocrine MoAs. Information on endocrine MoAs can be obtained from animal studies that include parameters measuring effects associated with endocrine MoAs. However, only certain standard test methods have been designed or updated so far to include these parameters (OECD TGs 407, 416, 421/422, 443, 206, 231, 229, 230, 234, 240, 241). Direct information on various mechanisms for endocrine disturbing properties (such as hormone receptor binding) can be obtained from *in vitro* methods developed for this purpose, however there is currently no *in vitro* method addressing thyroid disruption in mammals.

Methods that combine information on potential MoAs and adverse effects are animal studies, although the direct link between the MoAs and adverse effects may not be clear. None of the non-animal methods can currently reliably predict the link of endocrine mechanisms or MoAs with a certain adverse effect. Animal experiments are far less suitable to provide information on the MoA. The integrative nature of the animal model comes at the price of an experimental black box that provides, at best, some physiological details about the underlying mechanisms, but hardly any molecular understanding. However, any information on endocrine MoAs is valuable and adds to the overall evaluation on the mechanisms/MoAs of adverse effects. Information on MoAs can be used as supporting evidence in testing strategies.

The main challenges are to cover the relevant parts of the signalling network with the test system and the link to adverse effects. Under the current regulatory environment, it is also necessary to be able to specify the adverse effects (e.g. reproductive toxicity, carcinogenicity, etc.). A further challenge is how to address the potential long-term and cumulative effects, especially after low-dose exposure. Non-animal approaches are useful in predicting the endocrine activity and could be used for (de)prioritisation although many false positives may be predicted.

B.12 Bioaccumulation in fish

B.12.1 Description of the endpoint

Bioconcentration only refers to the accumulation by an aquatic organism of the substance dissolved in water. It is expressed as a bioconcentration factor, BCF.

Bioaccumulation is a complex ecological process which refers to the result of uptake of a substance in an organism from all environmental sources (including water, food and sediment) as well as metabolism and excretion. Highly bioaccumulative substances may transfer through the food web. It is expressed as a bioaccumulation factor, BAF.

Biomagnification refers to accumulation through the food chain. It is expressed as a biomagnification factor, BMF.

REACH Annex IX indicates that information on bioaccumulation in aquatic species, preferably fish, is required for substances manufactured or imported in quantities of 100 tonnes per year or more. Information on bioaccumulation of a substance is used for hazard classification, PBT assessment and for modelling exposure in the food chain for risk assessment. Although bioaccumulation is not a specified endpoint below 100 tonnes per year, information may still be relevant for substances manufactured or imported in amounts of 10 or more tonnes per year (REACH Annex XIII), to conclude PBT/vPvB assessment in the chemical safety report (CSR). The purpose of the PBT/vPvB assessment is to identify substances that are persistent, bioaccumulative and toxic (PBT), or very persistent and very bioaccumulative (vPvB). If a registrant cannot derive a definitive conclusion in the PBT/vPvB assessment using the relevant available information, he must, based on section 2.1 of REACH Annex XIII, generate the necessary information, regardless of his tonnage band (for further details, see the Guidance on IR&CSA – Chapter R.11). In such a case, the only possibility to refrain from testing or generating other necessary information is to treat the substance "*as if it is a PBT or vPvB*".

For the purpose of aquatic classification under the CLP Regulation, information on bioaccumulation (preferably a bioconcentration factor derived from an experimental study) is used in the absence of adequate long-term aquatic toxicity data, in conjunction with information on degradation.

The information requirements and the relevant test methods for the REACH Regulation, the BPR as well as the basis of the CLP criteria are presented in Section 12 of Appendix 3 to this document.

B.12.2 How to minimise vertebrate animal testing

B.12.2.1 Specific adaptation rules

When met, some of the specific adaptation conditions contained in Column 2 of REACH Annex IX or column 3 of BPR Annex II give a possibility not to perform new testing on vertebrate animals to fulfil the information requirements for bioaccumulation (for further details, see Section 12 of Appendix 3 to this document).

B.12.2.2 Replacement, reduction and refinement methods

The general rules for adaptation of REACH Annex XI and BPR Annex IV apply, such as grouping and read-across, QSARs and WoE adaptations. Guidance on the evaluation and validation of both animal and non-animal data can be found in Section R.7.10 of the Guidance on IR&CSA –

Chapter R.7c. Specific recommendations on the use of these data for PBT/vPvB assessment can be found in Section R.11.4.1.2 of the Guidance on IR&CSA – Chapter R.11.

For **grouping and read-across** supporting information may include hydrolysing and degradation properties, physicochemical properties such as water solubility, dissociation, volatility and partitioning constants (organic matter, lipids). Furthermore, the hypothesis should cover why bioaccumulation potential of the source substances is similar (or creates a trend) to that of the target substance. Bioaccumulation or uptake of the substance determines the concentration of the substance reaching the target sites of toxic action and thus affects the degree of toxic effects. Therefore, if bioaccumulation potential is expected to differ, also the effects are assumed to differ. In this case, a worst-case prediction of the effects of the target substance would be still acceptable and apply. For neutral organic substances log K_{ow} is often used as a surrogate parameter to describe the extent of the substance that may accumulate in an aquatic organism and cause the effects. General information on grouping and read-across is provided in Section A.1.4.

The following types of information are relevant for assessing the bioaccumulation potential and may be used as parts of **WoE adaptation**:

- existing aquatic bioconcentration or bioaccumulation studies;
- *in vitro* data on metabolism in combination with kinetics of uptake and depuration;
- read-across with structurally similar substances and QSAR approaches;
- terrestrial or benthic accumulation studies;
- field data concerning biomagnification and bioaccumulation;
- toxicokinetics information on laboratory mammals and humans;
- toxicokinetics in aquatic organisms, birds and mammals;
- detection of elevated levels in biota;
- physicochemical properties (e.g. molecular size, log K_{ow}, water solubility)
- uptake and absorption efficiency;
- absence/presence of chronic toxicity.

Further guidance on the WoE for bioaccumulation is available in the Guidance on IR&CSA – Chapter R.11 and Chapter R.7c.

Mammalian toxicokinetic studies may provide useful information in a WoE for bioaccumulation assessment. Further guidance is available in Section R.7.10.15 of the Guidance on IR&CSA – Chapter R.7c. Metrics to consider include:

- metabolic capacity/rate constants;
- affinity for lipid or blood-rich tissues, which could include the volume of distribution (VD);
- the time taken to reach a steady-state (plateau) concentration in tissues;
- uptake efficiency and clearance, and elimination rates/half-lives.

Regarding alternative species, REACH Annex IX, 9.3.2 specifies that bioaccumulation tests for aquatic species should preferably be conducted in **fish**. However this does not exclude the possibility to use alternative test species within WoE adaptation, e.g. bivalve molluscs (oyster) or sediment-dwelling benthic oligochaetes where this is justified. By definition, **benthic species** live in sediment rather than in the water column per se. However, sediments are part of the aquatic ecosystems, therefore bioaccumulation tests with benthic species are relevant for the aquatic compartment. As for tests performed with **bivalve molluscs**, it should be noted that bivalves tend to stop feeding in the presence of toxins. Therefore, it is important that bioaccumulation tests with bivalves should be performed at concentrations well below the short-term toxicity concentration. Guidance on this can be found in Section R.7.10 of the Guidance on IR&CSA – Chapter R.7c. Invertebrate bioaccumulation tests include ASTM E1022-94 which describes a method for measuring bioconcentration in saltwater bivalve molluscs

using aqueous exposure [222]. It is similar to OECD TG 305, with modifications for molluscs (such as size, handling and feeding regime).

These studies with alternative species are particularly relevant when a valid fish BCF is not available. The OECD TG 315 on bioaccumulation in sediment-dwelling benthic oligochaetes is the preferred method for generating additional information. When it can be demonstrated that sediment is the compartment of concern for a substance, the biota-sediment accumulation factor (BSAF) obtained from such a study can be used for risk assessment and can be used as part of the WoE. When accumulation from the porewater is expected to dominate, bioaccumulation could be expressed as a BCF between the organism and dissolved pore water concentrations. Invertebrate species may have a lower metabolic capacity than fish species, e.g. as is the case for polycyclic aromatic hydrocarbons [223]. Bioaccumulation in these invertebrates may therefore be higher than in fish under the same exposure.

Fish *in vitro* metabolism methods have the potential to provide important data for bioaccumulation assessments since the degree of metabolism affects the amount of bioaccumulation. Approaches for using *in vitro* data to determine metabolic capacity have been described and studied in several test systems. Two methods have been proposed as OECD TGs (see OECD Project 3.13 in OECD Work plan for the Test Guidelines Programme 2017 [22]). The draft test guidelines are "*Determination of in vitro intrinsic clearance using cryopreserved rainbow trout hepatocytes (RT-HEP)*" and "*Determination of in vitro intrinsic clearance using rainbow trout liver S9 sub-cellular fraction (RT-S9)*" (see drafts on OECD website).

Although *in vitro* data on fish metabolism is not a standard REACH or BPR information requirement, results of such studies can support the bioaccumulation assessment and can be considered as part of a WoE adaptation. Further information on *in vitro* methods is available in the Guidance on IR&CSA – Chapter R.7c and in the Draft OECD Guidance Document on "*Determination of in vitro intrinsic clearance using cryopreserved hepatocytes (RT-HEP) or liver S9 sub-cellular fractions (RT-S9) from rainbow trout and extrapolation to in vivo intrinsic clearance, 2017*" (see draft on OECD website and [224, 225]).

QSAR model predictions may be used to estimate bioaccumulation potential, when accompanied by sufficient documentation according to REACH Annex XI, 1.3. Further information is provided in Section A.1.4. Several free and commercial computational software programmes are available to predict the bioaccumulation in fish. Most of the available QSAR models are developed to mimic the correlation between log K_{ow} and bioaccumulation behaviour as closely as possible. For most neutral organic substances, empirical evidence shows that BCF typically increases with the lipophilicity/hydrophobicity of the chemicals up to a certain level and then gradually decreases. This behaviour is explained by the capacity of more lipophilic chemicals to permeate fish gills to a greater extent than hydrophilic chemicals. However, when the chemicals are too large in size or have too low solubility in water (above log K_{ow} values of five or six), their capacity to permeate the gills or their availability in the test system is decreased. Consequently bioaccumulation potential decreases.

Examples of freely available software for predicting bioaccumulation in fish are OECD QSAR Toolbox, BCFBAF in EPISuite, VEGA, CASE Ultra (MultiCASE) and T.E.S.T. BCFBAF and CATALOGIC (commercial software from OASIS) take into account metabolism and other mitigating parameters.

To avoid unnecessary animal testing, an **integrated testing strategy** for PBT/vPvB assessment should be followed where the potential for persistence should normally be assessed before bioaccumulation potential. When it is clear that the substance is persistent (P) or very persistent (vP), a stepwise approach should be followed to elucidate whether the bioaccumulative (B) or very bioaccumulative (vB) criterion is fulfilled (see the Guidance on IR&CSA – Chapter R.11). In this case, if screening information data (e.g. log K_{ow} and QSAR) show that there is a concern that the substance is likely to be B or vB, then further

bioaccumulation assessment is required using a WoE (see above). If it is not possible to reach a definitive conclusion on whether or not the substance is B/vB, the generation of new bioaccumulation information is required. The new information is usually data from experimental fish bioconcentration or bioaccumulation studies.

It should be also noted that even if not required for PBT/vPvB assessment, information on bioaccumulation potential may still be needed for environmental classification purposes, under the CLP Regulation.

As mentioned in Section 12.3 of Appendix 3 to this document, OECD TG 305 "*Bioaccumulation in Fish: Aqueous and Dietary Exposure*" offers the option of **a minimised aqueous exposure test** (paragraphs 83-88) and of testing only one test concentration (paragraphs 49-51), provided that certain conditions are met. Both options can reduce the number of fish used if the test substance meets the specific conditions described in the test guideline and can be considered as a form of reduction method.

B.12.3 Challenges related to the development and application of non-animal prediction methods

The accuracy (and related acceptability) of computational methods depends on the availability of high quality experimental data to "train" the models and on the understanding of the toxicological MoA to better predict the effects. One limitation of most prediction models for bioaccumulation can be the prediction of metabolism and its rate. Therefore, the prediction models not considering metabolism may not adequately estimate the bioaccumulative potential. However, *in vitro* methods have been developed to provide information on metabolism in fish and may be used in conjunction with the models to receive a better estimate of the bioaccumulation potential. In addition, the combination of different software (i.e. using different fragments, different descriptor and/or mathematical modelling approaches, different applicability domain) or the consideration of other information in a WoE adaptation may increase the confidence in the overall assessment.

Furthermore, most computational methods are based on log K_{ow} which is not driving the bioaccumulation potential of many but not all substances. Therefore, such models might not be suitable for the calculation of a BCF value for the following types of substances (Appendix R.7.10-3 of the Guidance on IR&CSA – Chapter R.7c):

- inorganic substances
- UVCBs (see Section R.11.4.2.2 of the Guidance on IR&CSA – Chapter R.11)
- ionisable substances
- surface active substances (surfactants)
- organic substances that do not partition to lipid.

B.12.4 Future perspectives

Promising results from initial development studies have been indicated using *Hyalella azteca* in test conditions similar to OECD TG 305 by Treu *et al*. [226]. The uptake and accumulation of several lipophilic substances were investigated and analysed for their tissue concentrations in *Hyalella azteca*. The depuration and uptake rates were used to generate BCF estimates. The resulting BCFs were similar to those obtained from OECD TG 305 fish studies with rainbow trout. This may offer a future non-vertebrate option for test guideline development and once validated, it may be a possible refinement for the fish bioaccumulation test.

Computational tools are under constant development. New bioaccumulation tests that are performed can be used in training sets of the prediction models, which further allow to expand the applicability domain and improve reliability of the predictions on bioaccumulation potential in the future.

B.12.5 Summary and conclusion on bioaccumulation

Information on bioaccumulation is relevant for classification and labelling as well as risk assessment and PBT assessment. It is not mandatory for all registration dossiers. Under the REACH Regulation, it is required for substances manufactured or imported in quantities of 100 tonnes per year or more (although substances with 10 tonnes per year or more need also a conclusion on PBT/vPvB). Under the BPR, for active substances, it is considered as an additional dataset, only required under certain conditions. Under the CLP Regulation, all REACH provisions on bioaccumulation apply, apart from the difference with the REACH Regulation in the BCF threshold values. A BCF of ≥500 is used to indicate a bioaccumulation potential under the CLP Regulation whereas a BCF of >2000/5000 is used in PBT/vPvB assessment.

There are many alternative approaches to the fish bioaccumulation test currently available. Read-across and grouping can be used alone, as well as prediction methods. Furthermore, invertebrate tests, *in vitro* methods (developed to provide information on metabolic capacity in fish), QSAR predictions, and read-across approaches can be used together with the other available data in WoE adaptations. Specific adaptation rules also apply when supported by clear justifications and evidence showing that the substance has low potential for bioaccumulation or cannot pass biological membranes, or that the exposure is unlikely to occur. Relevant developments are ongoing to reduce the need for vertebrate testing.

B.13 Fish toxicity

B.13.1 Description of the endpoint

Effects of substances on organisms living in the water are usually determined by testing on organisms representing the three trophic levels, i.e. plants (or algae), invertebrates (crustaceans such as *Daphnia* spp.) and vertebrates (fish) [10]. For the purpose of this report (avoiding the use of vertebrate animals), the focus is on fish. Fish are the most abundant aquatic vertebrates and are key prey and predators in aquatic food webs. This makes them an important test subject to represent higher trophic levels in aquatic systems.

The information on aquatic toxicity may be used for classification and labelling under the CLP Regulation, the derivation of predicted no-effect concentration (PNEC) values for use in risk assessment, and for the PBT assessment. In general, the lowest of the available toxicity values of the different trophic levels (fish, crustacean, algae or aquatic plants) are used to determine the hazard category (for CLP), derive the PNEC (for risk assessment) or determine if the substance fulfils the toxicity criterion (for PBT assessment). Furthermore, aquatic toxicity data are also used in combination with the (log) Koc[4] to predict the hazard to soil or sediment organisms when no experimental results with these specific organisms are available.

Fish are also used for the identification of substance-specific endocrine MoAs. Endocrine sensitive biomarkers like vitellogenin, secondary sex characteristics and sex ratio allow the identification of androgenic, oestrogenic or steroidogenetic MoAs. The OECD Conceptual Framework for endocrine disruptor testing provides information on various tests that can be used to identify endocrine disruption [227]. However, for identification of an endocrine MoA, a relevant fish test may be appropriate depending also on the concern. Ecotoxicological studies with parameters sensitive to endocrine MoA are shown in Section 13 of Appendix 3 to this document.

Generally, short-term aquatic toxicity to fish (also referred to as acute toxicity) is studied by exposing fish to the test substance for a period of 96 hours. Mortalities are recorded at 24, 48, 72 and 96 hours and the concentrations which kill 50 % of the fish (LC_{50}) are determined where possible.

Several long-term fish studies (also generally referred to as chronic toxicity) are available (see Section 13.3 of Appendix 3 to this document). Chronic toxicity testing should cover measurements of toxicity over sensitive life stages (e.g. exposure of fertilised eggs until the control fish reach a juvenile life stage). Lethal and sub-lethal effects are assessed and no observed effect concentration (NOEC) and/or ECx (e.g. EC_{10}, EC_{20}) are determined to estimate the concentration that would cause a x % change in the effect measured. Depending on the purpose of the study even a full generation or a multi-generation study may be needed, or a study where measurements to identify endocrine disruption in fish are made.

Information on short-term toxicity to fish under the REACH Regulation is required for substances manufactured in amounts of 10 tonnes per year or more, while substances manufactured in amounts of 100 tonnes per year or more require information on long-term toxicity to fish. Long-term testing is also needed for substances manufactured in amounts of 10 tonnes per year or more if the substance is poorly water soluble, or if the toxicity criterion needs to be further investigated for PBT assessment. More information on information requirements and the relevant test methods for the REACH Regulation, the BPR as well as the basis of the CLP criteria are presented in Section 13 of Appendix 3 to this document.

[4] Koc - The soil organic carbon-water partitioning coefficient is the ratio of the mass of a chemical that is adsorbed in the soil per unit mass of organic carbon in the soil per the equilibrium chemical concentration in solution.

B.13.2 How to minimise vertebrate animal testing

B.13.2.1 Specific adaptation rules

When met, some of the specific adaptation conditions contained in Column 2 of REACH Annexes VIII-IX or column 3 of BPR Annex II give a possibility not to perform new testing on vertebrate animals to fulfil the information requirements for short- or long-term fish toxicity (for further details, see Section 13 of Appendix 3 to this document).

B.13.2.2 Replacement, reduction and refinement methods for short-term and long-term fish toxicity tests

The general rules for adaptation of REACH Annex XI and BPR Annex IV apply, such as grouping and read-across and WoE adaptations.

Grouping and read-across is applicable also to this information requirement and has also been used under the BPR for groups of substances to avoid further testing (e.g. quaternary ammonium compounds).

Concluding on aquatic pelagic toxicity may be achieved by using **WoE adaptation**. The prerequisite is that the results of the WoE lead to a regulatory outcome equivalent to that obtained by standard testing. The three regulatory processes need to be considered: classification and labelling, PBT assessment and environmental risk assessment.

The REACH Guidance on IR&CSA – Chapter R.7b, and especially Section *R.7.8.5 Conclusions for aquatic pelagic toxicity and integrated testing strategy (ITS)*, outlines a systematic approach of how to use all available data in a WoE adaptation. It provides a step-wise procedure for the assessment of different types of information, which might be helpful to come to an overall conclusion. The scheme proposes a flexible sequence of steps (characterisation of the substance, analysis of the mode of action, evaluation of existing *in vivo* and *in vitro* data, data from analogous QSAR model predictions), the order of which depends on the quality and quantity of data and might be changed.

Even though not mentioned in the Section *R.7.8.5 Conclusions for aquatic pelagic toxicity and integrated testing strategy (ITS)*, the fish embryo acute toxicity (FET) test (OECD TG 236) can be used within a WoE together with other independent, adequate, relevant and reliable sources of information.

The potential of the FET to fulfil the standard information requirement for the acute fish toxicity test was studied in an ECHA commissioned study entitled "*Analysis of the relevance and adequateness of using Fish Embryo Acute Toxicity test (FET) Test Guideline (OECD TG 236) to fulfil the information requirements and addressing concerns under REACH*" (May 2015). The Member State Committee was also consulted on the report. The project aimed to gather and analyse publicly available data on FET, comparing it with available data on standard acute fish toxicity (AFT) to investigate the predictive power of the FET test. Furthermore, the study aimed at defining the applicability domain of the FET test (e.g. chemical structure, MoA and several key physicochemical characteristics).

In light of the analysis made by ECHA, there are still certain limitations in the use of the FET test and therefore it should only be used within a WoE adaptation. If used, the limitations identified in the ECHA analysis need to be taken into account. The conclusions of the scientific analysis performed within ECHA's project are available in the report prepared by the consultant at: http://echa.europa.eu/web/guest/publications/technical-scientific-reports.

Regarding CLP, a WoE determination will be required where there are no sufficient data that can be used directly for comparison with the CLP criteria. Further information is provided in Section A.1.3.2.

Several free and commercial **computational software methods** are available to predict aquatic toxicity. They include models to predict both short and/or long-term aquatic toxicity for all trophic levels for organic monoconstituent substances. Most of these approaches rely on at least one of the two following assumptions: 1) that the MoA of a substance will depend on its chemical class (i.e. the functional groups in its structure); and 2) that the aquatic toxicity of a substance will be proportional to its bioaccumulation potential (described by lipophilicity in the models). Most non-reactive, non-ionisable organic substances are expected to exert their toxicity with a simple non-polar narcosis mechanism ("*baseline toxicity*"). The toxicity of this class of substances is usually predicted with good accuracy, while toxicity of substances having other modes of actions are less well predicted (see Section B.13.3 below).

Examples of freely available software for predicting aquatic toxicity are OECD QSAR Toolbox, ECOSAR, VEGA, T.E.S.T. and Danish QSAR DB. See also ECHA's Practical guide on "*How to use and report (Q)SARs*" and illustrative examples for short and long-term fish toxicity. Further information is provided in Section A.1.2.1.

If a new acute fish toxicity test needs to be conducted, there is a possibility to reduce the number of fish tested with a **limit test**. A limit test may be performed at 100 mg (active ingredient)/l to demonstrate that the LC_{50} is greater than this concentration. Binomial theory dictates that when 10 fish are used with zero mortality, there is a 99.9 % confidence that the LC_{50} is greater than 100 mg/l. With seven, eight or nine fish, the absence of mortality provides at least 99 % confidence that the LC_{50} is greater than the concentration used in the limit test. If any mortalities occur, a full study should be conducted.

The **threshold approach** offers a possibility to reduce the number of fish to be used in acute aquatic toxicity testing when a test on fish is required. It takes into consideration that only the lowest value of the acute toxicity in species of three trophic levels is considered for regulatory purposes. The approach is described in the OECD GD 126 "*Short Guidance on the threshold approach for acute fish toxicity*" [228].

The threshold approach addresses fish toxicity by initially using a single-concentration test (limit test) requiring less fish compared to the full acute fish toxicity study. The selection of a single concentration is based on the derivation of a threshold concentration (TC) from reliable algae and acute invertebrate (e.g. *Daphnia*) toxicity data. Fish toxicity is then tested at the TC to consider if fish are more or less sensitive than groups/species for which an E/LC_{50} is available. If no mortality occurs in the limit test using the TC, the TC might be used as a surrogate of the LC_{50} value in the further hazard or risk assessment.

In the context of the PBT/vPvB assessment, using **integrated testing strategy**, a conclusion on the P and B properties should normally be drawn before further T-testing is considered. If the substance is found to be both persistent (P) and bioaccumulative (B) then a chronic toxicity study is required (except if the substance meets the criteria for classification for carcinogenicity, mutagenicity, reproductive toxicity or for chronic toxicity according to the CLP regulation (see section 1.1.3, points (b) and (c) of Annex XIII to the REACH Regulation). Normally, the testing sequence for a conclusion on T based on chronic data in *Daphnia* and then fish, unless there are indications that fish is the most sensitive group. If the T-criterion is fulfilled by the chronic algae or *Daphnia* data, a chronic fish test is not necessary and should therefore not be carried out to avoid unnecessary testing on vertebrate animals.

If long-term toxicity testing is triggered by the CSA (Annex IX, 9.1.6), testing fish may still be avoided. According to the integrated testing strategy presented in Guidance on IR&CSA – Chapter R7b (Section R.7.8.5, including Figure R.7.8-4), if based on acute aquatic toxicity data neither fish nor invertebrates are shown to be substantially more sensitive, long-term studies

may be required on both. In such case, according to the ITS, the *Daphnia* study is to be conducted first. If based on the results of the long-term *Daphnia* study and the application of a relevant assessment factor, no risks are observed (PEC/PNEC<1), no long-term fish testing may need to be conducted. However, if a risk is indicated, the long-term fish study needs to be conducted.

B.13.3 Challenges related to the development and application of non-animal prediction methods

The accuracy (and related acceptability) of computational methods depends on the availability of high quality experimental data to "train" the models and on the understanding of the toxicological MoA to better predict the effects. In the specific case of fish toxicity, the following challenges are related to the training sets:

- the possibility to use different fish species for performing the tests – the inter-species differences in sensitivity increase the uncertainty/variability of the results. This is true especially for long-term toxicity tests, therefore making the development of reliable models more challenging;
- data on long-term toxicity is scarce – the training sets for long-term fish toxicity models are therefore not large, which reduces the applicability domain in terms of chemical and mechanistic space;
- chemicals which act with MoAs more complex than the simple narcosis are less common in the training sets – there is limited knowledge of their MoA and consequently they are more difficult to predict. For example, substances with reactive functional groups are predicted to have higher toxicity compared to the baseline, but an accurate prediction of the exact value is often difficult to obtain. The chemical class of neutral organics represent the widest class of substances and therefore the models are well trained to predict their toxicity.

In many cases, it is recommended to support a computational result with other information, as a form of WoE. Furthermore, each QSAR prediction should be accompanied by scientific justification and documentation according to REACH Annex XI, 1.3. Similarly, under the CLP Regulation, relevant information and suitable assumptions used for a QSAR model could allow for QSAR derived toxicity data to be considered under a WoE determination.

The validity and interpretation of the results from prediction models always depends on the assumptions and data/training sets used in the model. As a result, these assumptions and data must be relevant to the question being asked to render the outcome of use in a regulatory context.

B.13.4 Future perspectives

Considering the current work on AOPs in the field of aquatic toxicity and the efforts in reducing the use of studies on juvenile/adult fish, discussions are still ongoing on how to incorporate this in regulatory science. Furthermore, computational methods are constantly being developed and new fish toxicity data is becoming available, increasing the chances for 3Rs approaches to be accepted in the regulatory context.

Regarding FET, in May 2016 ECHA hosted an expert Workshop on the potential regulatory application of the fish embryo acute toxicity (FET) test under the REACH and CLP regulations and the BPR. The workshop was jointly organised by ECHA and the German Environment Agency (Umweltbundesamt, UBA), with support of UBA Austria and EURL ECVAM (European Commission DG Joint Research Centre, Directorate F, unit F.3) in the steering committee. The aim of the workshop was to exchange views on the potential regulatory application of the FET and explore possibilities on how the FET might be used as a part of WoE approaches in the EU regulatory context (REACH, Biocides and CLP) to adapt standard information requirements for

acute fish toxicity. During the workshop, the research needs and areas for further developments to improve usability of FET for regulatory purposes were identified. Moreover, industry was invited to include available FET data in the WoE approach(es) in their registrations to gain experience and to build the case studies that might be used as best practice examples.

In addition, at OECD level, there are currently ongoing discussions on how to integrate the FET into the OECD GD 126 on the threshold approach for acute fish toxicity (OECD project 2.54) [228].

B.13.5 Summary and conclusion on fish toxicity

Aquatic toxicity studies are important information sources for assessing the hazard and risk to freshwater and marine organisms living in the water column on exposure to a substance.

Under the REACH Regulation, information on the short-term toxicity to fish is required when a substance is registered in quantities higher than 10 tonnes per year. Information on long-term toxicity to fish is normally required when a substance is registered in quantities higher than 100 tonnes per year (or 10 tonnes per year if the substance is poorly water soluble or if the toxicity criterion needs to be further investigated for PBT assessment). Under the BPR for active substances, short-term toxicity testing on fish is part of the core dataset and must always be provided. It is nevertheless specified that when short-term fish toxicity data are required, the threshold approach (tiered strategy) [228] should be applied and that the study does not need to be conducted if a valid long-term aquatic toxicity study on fish is available. Long-term toxicity testing on fish is an additional data requirement only for certain product-types (where for example continuous release to the aquatic compartment occurs).

General adaptation rules such as QSAR, read-across and grouping, and WoE allow using methods to replace or reduced vertebrate testing alone or together if adequately justified and if the uncertainties have been considered. Furthermore, new approaches are being developed to predict fish toxicity (e.g. toxicokinetic modelling approaches combined with *in vitro* toxicity tests, integration of the FET into the OECD GD 126) which may further reduce animal testing in the future.

References

[1] S. Adler, D. Basketter, S. Creton, O. Pelkonen, J. van Benthem, V. Zuang, K. E. Andersen, A. Angers-Loustau, A. Aptula, A. Bal-Price, E. Benfenati, U. Bernauer, J. Bessems, F. Y. B. A. Bois, E. Brandon, S. Bremer, T. Broschard, S. Casati, S. Coecke, R. Corvi, M. Cronin, G. Daston, W. Dekant, S. Felter, E. Grignard, U. Gundert-Remy, T. Heinonen, I. Kimber, J. Kleinjans, H. Komulainen, R. Kreiling, J. Kreysa, S. Batista Leite, G. Loizou, G. Maxwell, P. Mazzatorta, S. Munn, Phrakonkham, P. Phrakonkham, A. Piersma, A. Poth, Repetto, G. Repetto, V. Rogiers, G. Schoeters, M. Schwarz, R. Serafimova, H. Tähti, E. Testai, J. van Delft, H. van Loveren, M. Vinken, A. Worth and M. Zaldivar, "Alternative (non-animal) methods for cosmetics testing: current status and future prospects—2010," 2011. [Online]. Available: https://eurl-ecvam.jrc.ec.europa.eu/eurl-ecvam-status-reports/files/article-status-report-2010.

[2] ECHA, "ECHA report on The Use of Alternatives to Testing on Animals for the REACH Regulation, third report under Article 117(3) of the REACH Regulation," 2017. [Online]. Available: https://echa.europa.eu/documents/10162/13639/alternatives_test_animals_2017_en.pdf/075c690d-054c-693a-c921-f8cd8acbe9c3.

[3] W. M. S. Russell and R. L. Burch, "The Principles of Humane Experimental Technique," 1959. [Online]. Available: http://altweb.jhsph.edu/pubs/books/humane_exp/het-toc.

[4] M. Balls, "Replacement of animal procedures: alternative in research, education and testing," *Laboratory Animals,* no. 28, pp. 193-211, 1994.

[5] OECD, "Revised Guidance Document on Developing and Assessing Adverse Outcome Pathways. Series on Testing and Assessment No. 184 (ENV/JM/MONO(2013)6)," 2017. [Online]. Available: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono%282013%296&doclanguage=en.

[6] C. MacKay, M. Davies, V. Summerfield and G. Maxwell, "From Pathways to People: Applying the Adverse Outcome Pathway (AOP) for Skin Sensitisation to Risk Assessment.," *ALTEX,* vol. 30, pp. 473-86 , 2013.

[7] C. Wittwehr, H. Aladjov, G. Ankley, H. Byrne, J. de Knecht, E. Heinzle, G. Klambauer, B. Landesmann, M. Luijten, C. MacKay, G. Maxwell, M. Meek, A. Paini, E. Perkins, T. Sobanski, D. Villeneuve, K. Waters and M. Whelan, "How Adverse Outcome Pathways Can Aid the Development and Use of Computational Prediction Models for Regulatory Toxicology.," *Toxicol Sci,* vol. 155, no. 2, pp. 326-36, 2017.

[8] R. Serafimova, M. Fuart Gatnik and A. Worth, "Review of QSAR models and software tools for predicting genotoxicity and carcinogenicity. JRC Technical Report EUR 24427 EN.," Publications Office of the European Union, Luxembourg, 2010.

[9] A. Worth, S. Lapenna, E. Lo Piparo, A. Mostrag-Szlichtyng and S. R, "The Applicability of Software Tools for Genotoxicity and Carcinogenicity Prediction: Case Studies relevant to the Assessment of Pesticides. JRC Scientific and Technical Report EUR 24640 EN - 2010," Publications Office of the European Union, Luxemburg, 2010.

[10] A. Worth, J. Barroso, S. Bremer, J. Burton, S. Casati, S. Coecke, R. Corvi, B. Desprez, C. Dumont, V. Gouliarmou, M. Goumenou, R. Gräpel, C. Griesinger, M. Halder, A. Janusch Roi, A. Kienzler, F. Madia, S. Munn, M. Nepelska, A. Paini, A. Price, P. Prieto, A. Rolaki, M. Schäffer, J. Triebe, M. Whelan, C. Wittwehr and V. Zuang, "EURL ECVAM Science and policy report on Alternative methods for regulatory toxicology - a state-of-the-art review (EUR 26797 EN)," JRC Science and policy report , 2014.

[11] D. Asturiol and A. Worth, "The Use of Chemical Reactivity Assays in Toxicity Prediction. JRC Scientific and Technical Report EUR 24870 EN - 2011," Publications Office of the European Union, Luxembourg, 2011.

[12] J. Liu, K. Mansouri, R. S. Judson, M. T. Martin, H. Hong, M. Chen, X. Xu, R. S. Thomas

and I. Shah, "Predicting hepatotoxicity using ToxCast in vitro bioactivity and chemical structure," *Chem Res Toxicol,* vol. 28, no. 4, pp. 738-51, 2015.

[13] C. W. Scott, M. F. Peters and Y. M. Dragan, "Human induced pluripotent stem cells and their use in drug discovery for toxicity testing," *Toxicol Lett,* vol. 219, no. 1, pp. 49-58, 2013.

[14] R. L. 3. Gieseck, L. Vallier and N. R. Hannan, "Generation of Hepatocytes from Pluripotent Stem Cells for Drug Screening and Developmental Modeling," *Methods Mol Biol,* vol. 1250, pp. 123-42, 2015.

[15] L. Suter-Dick, P. M. Alves, B. J. Blaauboer, K. D. Bremm, C. Brito, S. Coecke, B. Flick, P. Fowler, J. Hescheler, M. Ingelman-Sundberg, P. Jennings, J. M. Kelm, I. Manou, P. Mistry, A. Moretto, A. Roth, D. Stedman, B. van de Water and M. Beilmann, "Stem cell-derived systems in toxicology assessment," *Stem Cells Dev,* vol. 24, no. 11, pp. 1284-96, 2015.

[16] A. K. Sjogren, M. Liljevald, B. Glinghammar, J. Sagemark, X. Q. Li, A. Jonebring, I. Cotgreave, G. Brolén and T. B. Andersson, "Critical differences in toxicity mechanisms in induced pluripotent stem cell-derived hepatocytes, hepatic cell lines and primary hepatocytes," *Arch Toxicol,* vol. 88, no. 7, pp. 1427-1437, 2014.

[17] A. C. Romero, E. Del Río, E. Vilanova and M. A. Sogorb, "RNA transcripts for the quantification of differentiation allow marked improvements in the performance of embryonic stem cell test (EST)," *Toxicol. Lett.,* vol. 238, no. 3, pp. 60-69, 2015.

[18] P. T. Theunissen, J. F. Robinson, J. L. Pennings, M. H. van Herwijnen, J. C. Kleinjans and A. H. Piersma, "Compound-specific effects of diverse neurodevelopmental toxicants on global gene expression in the neural embryonic stem cell test (ESTn)," *Toxicol Appl Pharmacol,* vol. 262, no. 3, pp. 330-40, 2012.

[19] E. C. Tonk, J. F. Robinson, A. Verhoef, P. T. Theunissen, J. L. Pennings and A. H. Piersma, "Valproic acid-induced gene expression responses in rat whole embryo culture and comparison across in vitro developmental and non-developmental models," *Reprod Toxicol,* vol. 41, pp. 57-66, 2013.

[20] A. R. Perestrelo, A. C. Águas, A. Rainer and G. Forte, "Microfluidic Organ/Body-on-a-Chip Devices at the Convergence of Biology and Microengineering.," *Sensors (Basel),* vol. 15, no. 12, pp. 31142-70, 2015.

[21] A. Balijepalli and V. Sivaramakrishan, "Organs-on-chips: research and commercial perspectives.," *Drug Discov Today,* vol. 22, no. 2, pp. 397-403, 2017.

[22] OECD, "OECD Work Plan for the Test Guidelines Programme," 2017. [Online]. Available: http://www.oecd.org/chemicalsafety/testing/TGP%20work%20plan_August%202017.pdf.

[23] M. Yoon, B. J. Blaauboer and H. J. Clewell, "Quantitative in vitro to in vivo extrapolation (QIVIVE): An essential element for in vitro-based risk assessment," 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0300483X15000372.

[24] N. I. Kramer, E. Di Consiglio, B. J. Blaauboer and E. Testai, "Biokinetics in repeated-dosing in vitro drug toxicity studies," 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0887233315002180?via%3Dihub.

[25] OECD, "Series on Testing and Assessment No. 238, Overview of the set of OECD Genetic Toxicology Test Guidelines and updates performed in 2014-2015 (ENV/JM/MONO(2016)33)," 2016. [Online]. Available: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono%282016%2933&doclanguage=en.

[26] ECHA, "New Approach Methodologies. Proceedings of a scientific workshop," 2016. [Online]. Available: https://echa.europa.eu/documents/10162/22816069/scientific_ws_proceedings_en.pdf/a2087434-0407-4705-9057-95d9c2c2cc57.

[27] B. J. Blaauboer, K. Boekelheide, H. J. Clewell, M. Daneshian, M. M. L. Dingemans, A. M. Goldberg, M. Heneweer, J. Jaworska, N. I. Kramer, M. Leist, H. Seibert, E. Testai, R. J. Vandebriel, J. D. Yager and J. Zurlo, "The use of biomarkers of toxicity for integrating in vitro hazard estimates into risk assessment for humans," *ALTEX,* vol. 29, no. 4, pp. 411-25, 2012.

[28] B. J. Blaauboer, A. R. Boobis, B. Bradford, A. Cockburn, A. Constable, M. Daneshian, G. Edwards, J. A. Garthoff, B. Jeffery, C. Krul and J. Schuermans, "Considering new methodologies in strategies for safety assessment of foods and food ingredients.," *Food Chem Toxicol,* vol. 91, pp. 19-35, 2016.

[29] OECD, "Series on Testing and Assessment No. 255, Guidance Document on the Reporting of Defined Approaches to be Used Within Integrated Approaches to Testing and Assessment (ENV/JM/MONO(2016)28)," 2016. [Online]. Available: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono(2016)28&doclanguage=en.

[30] OECD, "Series on Testing and Assessment No. 34, Guidance document on the validation and international acceptance of new or updated test methods for hazard assessment (ENV/JM/MONO(2005)14)," 2005. [Online]. Available: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono%282005%2914&doclanguage=en.

[31] R. S. Settivari, N. Ball, L. Murphy, R. Rasoulpour, D. R. Boverhof and E. W. Carney, "Predicting the future: opportunities and challenges for the chemical industry to apply 21st-century toxicity testing," *J Am Assoc Lab Anim Sci,* vol. 54, no. 2, pp. 214-23., 2015.

[32] OECD, "Series on Testing and Assessment No. 69, Guidance Document on the Validation of (Quantitative) Structure-Activity Relationship [(Q)SAR] Models (ENV/JM/MONO(2007)2)," 2007. [Online]. Available: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?doclanguage=en&cote=env/jm/mono%282007%292.

[33] OECD, "Series on Testing and Assessment No. 194, Guidance on Grouping of Chemicals, second edition (ENV/JM/MONO(2014)4)," 2014. [Online]. Available: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono%282014%294&doclanguage=en.

[34] OECD, "Series on Testing and Assessment No. 211, Guidance Document for Describing Non-Guideline In Vitro Test Methods (ENV/JM/MONO(2014)35)," 2014. [Online]. Available: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=ENV/JM/MONO%282014%2935&doclanguage=en.

[35] OECD, "Series on Testing and Assessment No. 203, Guidance Document proposes an integrated approach on testing and assessment (IATA) for skin corrosion and irritation (ENV/JM/MONO(2014)19)," 2014. [Online]. Available: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono(2014)19&doclanguage=en.

[36] OECD, " Series on Testing & Assessment No. 263, Guidance Document on an Integrated Approach on Testing and Assessment (IATA) for Serious Eye Damage and Eye Irritation (ENV/JM/MONO(2017)15)," 2017. [Online]. Available: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=ENV/JM/MONO(2017)15&doclanguage=en.

[37] OECD, "Series on Testing & Assessment No. 256, Guidance Document On The Reporting Of Defined Approaches And Individual Information Sources To Be Used Within Integrated Approaches To Testing And Assessment (IATA) For Skin Sensitisation (ENV/JM/MONO(2016)29)," 2016. [Online]. Available: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono(2016)29&doclanguage=en.

[38] SCENIHR (Scientific Committee on Emerging and Newly Identified Health Risks), "SCENIHR Memorandum on the use of the scientific literature for human health risk assessment purposes – weighing of evidence and expression of uncertainty, 19 March, 2012," 2012. [Online]. Available: http://ec.europa.eu/health/scientific_committees/emerging/docs/scenihr_s_001.pdf.

[39] A. Hardy, D. Benford, T. Halldorsson, M. Jeger, K. Knutsen, S. More, H. Naegeli, H. Noteborn, C. Ockleford, A. Ricci, G. Rychen, J. Schlatter, V. Silano, R. Solecki, D. Turck, E. Benfenati, Q. Chaudhry, P. Craig, G. Frampton, M. Greiner, A. Hart, C. Hogstrand, C. Lambre, R. Luttik, D. Makowski, A. Siani, H. Wahlstroem, J. Aguilera, J.-L. Dorne, A. Fernandez Dumont, M. Hempen, S. Valtueña Martínez, L. Martino, C. Smeraldi, A. Terron, N. Georgiadis and M. Younes, "Guidance on the use of the weight of evidence approach in scientific assessments," 2017. [Online]. Available: https://www.efsa.europa.eu/en/efsajournal/pub/4971.

[40] US OSHA (Occupational Safety and Health Administration), "(Draft) Guidance on Data Evaluation for Weight of Evidence Determination (OSHA-2016-0004)," 2016. [Online]. Available: https://www.osha.gov/weightofevidence/woe_guidance.pdf.

[41] B. van Ravenzwaay, S. Sperber, O. Lemke, E. Fabian, F. Faulhammer, H. Kamp, W. Mellert, V. Strauss, A. Strigun, E. Peter, M. Spitzer and W. T, "Metabolomics as read-across tool: A case study with phenoxy herbicides.," *Regul Toxicol Pharmacol,* vol. 81, pp. 288-304, 2016.

[42] ECHA, "ECHA Read-Across Assessment Framework (RAAF)," 2015. [Online]. Available: https://echa.europa.eu/documents/10162/13628/raaf_en.pdf.

[43] T. W. Schultz and M. T. D. Cronin, "Lessons learned from read-across case studies for repeated-dose toxicity," *Regul Toxicol Pharmacol,* vol. 88, pp. 185-91, 2017.

[44] ECHA, "Read-Across Assessment Framework (RAAF) - Considerations on multi-constituent substances and UVCBs," 2016. [Online]. Available: https://echa.europa.eu/documents/10162/13630/raaf_uvcb_report_en.pdf/3f79684d-07a5-e439-16c3-d2c8da96a316.

[45] J. S. MacDonald and R. T. Robertson, "Toxicity testing in the 21st century: a view from the pharmaceutical industry," *Toxicological Sciences,* vol. 110, no. 1, pp. 40-46, 2009.

[46] National Research Council; Division on Earth and Life Studies; Board on Environmental Studies and Toxicology; Institute for Laboratory Animal Research; Committee on Toxicity Testing and Assessment of Environmental Agents , "Toxicity Testing in the 21st Century: A Vision and a Strategy," The National Academies Press, Washington, D.C., 2007.

[47] H. Olson, G. Betton, D. Robinson, K. Thomas, A. Monro, G. Kolaja, P. Lilly, J. Sanders, G. Sipes, W. Bracken, M. Dorato, K. Van Deun, P. Smith, B. Berger and A. Heller, "Concordance of the toxicity of pharmaceuticals in humans and in animals," *Regul Toxicol Pharmacol,* vol. 32, no. 1, pp. 56-67, 2000.

[48] P. Browne, R. S. Judson, W. M. Casey, N. C. Kleinstreuer and R. S. Thomas, "Screening Chemicals for Estrogen Receptor Bioactivity Using a Computational Model," *Environ Sci Technol,* vol. 49, no. 14, pp. 8804-14, 2015.

[49] M. J. Schiffelers, B. Blaauboer, C. Hendriksen and W. Bakker, "Regulatory acceptance and use of 3R models: a multilevel perspective," no. 3, pp. 287-300, 2012.

[50] M. J. Schiffelers, B. J. Blaauboer, W. E. Bakker, S. Beken, C. F. Hendriksen, H. B. Koëter and C. Krul, "Regulatory acceptance and use of 3R models for pharmaceuticals and chemicals: expert opinions on the state of affairs and the way forward," *Regul Toxicol Pharmacol,* vol. 69, pp. 41-8, 2014.

[51] F. Busquet and T. Hartung, "The need for strategic development of safety sciences," *ALTEX,* vol. 34, no. 1, pp. 3-21, 2017.

[52] U. G. Sauer, E. H. Hill, R. D. Curren, H. A. Raabe, S. N. Kolle, W. Teubner, A. Mehling and R. Landsiedel, "Local tolerance testing under REACH: Accepted non-animal methods are not on equal footing with animal tests," *Altern Lab Anim,* vol. 44, no. 3,

pp. 281-99, 2016.

[53] T. Tralau, M. Oelgeschläger, R. Gürtler, G. Heinemeyer, M. Herzler, T. Höfer, H. Itter, T. Kuhl, N. Lange, N. Lorenz, C. Müller-Graf, U. Pabel, R. Pirow, V. Ritz, H. Schafft, H. Schneider, T. Schulz, D. Schumacher, S. Zellmer, G. Fleur-Böl, M. Greiner, M. Lahrssen-Wiederholt , A. Lampen, A. Luch, G. Schönfelder, R. Solecki, R. Wittkowski and A. Hensel, "Regulatory toxicology in the twenty-first century: challenges, perspectives and possible solutions," *Arch Toxicol,* vol. 89, no. 6, pp. 823-50, 2015.

[54] K. Rasmussen, M. González, P. Kearns, J. R. Sintes, F. Rossi and P. Sayre, "Review of achievements of the OECD Working Party on Manufactured Nanomaterials' Testing and Assessment Programme. From exploratory testing to test guidelines," *Regul Toxicol Pharmacol,* vol. 74, pp. 147-60, 2016.

[55] European Commission / SCCS, "The SCCS notes of guidance for the testing of cosmetic ingredients and their safety evaluation, 9th revision," 2015. [Online]. Available: http://ec.europa.eu/health/scientific_committees/consumer_safety/docs/sccs_o_190.pdf.

[56] A. Worth, K. Aschberger, D. Asturiol Bofill, J. Bessems, K. Gerloff, R. Graepel, E. Joossens, L. Lamon, T. Palosaari and A. Richarz, "Evaluation of the availability and applicability of computational approaches in the safety assessment of nanomaterials: Final report of the Nanocomput project, JRC Technical Reports (EUR 28617 EN)," Publications Office of the European Union, Luxembourg, 2017.

[57] EFSA (European Food Safety Authority), "Guidance for risk assessment of engineered nanomaterials," *EFSA Journal,* vol. 9, no. 5, p. 2140 [36 pp.], 2011.

[58] European Commission / SCCS (Scientific Committee on Consumer Safety), "Guidance on the safety assessment of nanomaterials in cosmetics (SCCS/1484/12)," 2012. [Online]. Available: http://ec.europa.eu/health/scientific_committees/consumer_safety/docs/sccs_s_005.pdf.

[59] M. Paparella, A. Colacci and M. N. Jacobs, "Uncertainties of Testing Methods: What Do We (Want to) Know About Carcinogenicity?," *ALTEX,* vol. 34, no. 2, pp. 235-252, 2017.

[60] N. Downes and J. Foster, "Regulatory Forum Opinion Piece: Carcinogen Risk Assessment: The Move from Screens to Science," *Toxicol Pharmacol,* vol. 43, pp. 1064-73, 2015.

[61] M. Paparella, A. Colacci and M. N. Jacobs, "Uncertainties of testing methods: What do we (want to) know about carcinogenicity?," *ALTEX,* 2016.

[62] E. Perkins, P. Antczak, L. Burgoon, F. Falciani, N. Garcia-Reyero, S. Gutsell, G. Hodges, A. Kienzler, D. Knapen, M. McBride and C. Willett, "Adverse Outcome Pathways for Regulatory Applications: Examination of Four Case Studies With Different Degrees of Completeness and Scientific Confidence.," *Toxicol Sci,* vol. 148, pp. 14-25, 2015.

[63] EFSA / EFSA Scientific Committee, "Guidance on Uncertainty in EFSA Scientific Assessment - draft," [Online]. Available: https://www.efsa.europa.eu/sites/default/files/160321DraftGDUncertaintyInScientificAssessment.pdf.

[64] EFSA, "Workshop report on integrated approach for testing and assessment of developmental neurotoxicity," EFSA Supporting publication 2017:EN-1191, 2017.

[65] H. J. Klimisch, M. Andreae and U. Tillmann, "A systematic approach for evaluating the quality of experimental toxicological and ecotoxicological data," *Regul Toxicol Pharmacol,* vol. 25, no. 1, pp. 1-5, 1997.

[66] D. Pamies, A. Bal-Price, A. Simeonov, D. Tagle, D. Allen, D. Gerhold, D. Yin, F. Pistollato, T. Inutsuka, K. Sullivan , G. Stacey, H. Salem, M. Leist, M. Daneshian, M. Vemuri, R. McFarland, S. Coecke, S. Fitzpatrick , U. Lakshmipathy, A. Mack, W. Wang, D. Yamazaki, Y. Sekino , Y. Kanda, L. Smirnova and T. Hartung, "Good Cell Culture Practice for stem cells and stem-cell-derived models.," *ALTEX,* vol. 34, no. 1, pp. 95-132, 2017.

[67] R. Kase, M. Korkaric, I. Werner and M. Ågerstrand, "Criteria for Reporting and Evaluating ecotoxicity Data (CRED): comparison and perception of the Klimisch and CRED methods for evaluating reliability and relevance of ecotoxicity studies," *Environ Sci Eur,* vol. 28, no. 1, p. 7, 2016.

[68] J. Kaltenhäuser, C. Kneuer, P. Marx-Stoelting, L. Niemann, J. Schubert, Stein, B and R. Solecki, "Relevance and reliability of experimental data in human health risk," *Regulatory Toxicology and Pharmacology,* no. 88, pp. 227-237, 2017.

[69] OECD, "Series on Principles of Good Laboratory Practice (GLP) and Compliance Monitoring No. 1, OECD Principles on Good Laboratory Practice (ENV/MC/CHEM(98)17)," 1998. [Online]. Available: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/mc/chem%2898%2917&doclanguage=en.

[70] OECD, "Series on Testing and Assessment No. 211, Guidance Document for describing non-guideline in vitro test methods (ENV/JM/MONO (2014)35)," 2014. [Online]. Available: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=ENV/JM/MONO%282014%2935&doclanguage=en.

[71] OECD, "Draft Guidance Document on Good In Vitro Method Practices (GIVIMP) for the development and implementation of in vitro methods for regulatory use in human safety assessment," 2017. [Online]. Available: http://www.oecd.org/env/ehs/testing/OECD%20Draft%20GIVIMP_v05%20-%20clean.pdf.

[72] D. Basketter, A. Poole and K. I, "Behaviour of chemical respiratory allergens in novel predictive methods for skin sensitisation," 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0273230017300600.

[73] R. Leggert, E. Ansoborlo, M. Bailey, D. Gregoratto, F. Paquet and D. Taylor, "Biokinetic data and models for occupational intake of lanthanoids," 2014. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/24475737.

[74] B. J. Blaauboer, "The necessity of biokinetic information in the interpretation of in vitro toxicity data," 2002. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/12513656.

[75] J. G. Bessems, G. Loizou, K. Krishnan, H. J. 3. Clewell, C. Bernasconi, F. Bois, S. Coecke, E. M. Collnot, W. Diembeck, L. R. Farcal, L. Geraets, U. Gundert-Remy, N. Kramer, G. Küsters, S. B. Leite, O. R. Pelkonen, K. Schröder, E. Testai, I. Wilk-Zasadna and J. M. Zaldívar-Comenges, "PBTK modelling platforms and parameter estimation tools to enable animal-free risk assessment: recommendations from a joint EPAA-EURL ECVAM ADME workshop," *Regul Toxicol Pharmacol,* vol. 68, no. 1, pp. 119-39, 2014.

[76] S. A. Saghir, M. J. Bartels, D. L. Rick, A. T. McCoy, R. J. Rasoulpour, R. G. Ellis-Hutchings, M. Sue Marty, C. Terry, J. P. Bailey, R. Billington and J. S. Bus, "Assessment of diurnal systemic dose of agrochemicals in regulatory toxicity testing-an integrated approach without additional animal use," *Regul Toxicol Pharmacol,* vol. 63, no. 2, pp. 321-32, 2012.

[77] J. E. Sager, J. Yu, I. Ragueneau-Majlessi and I. N, "Physiologically Based Pharmacokinetic (PBPK) Modeling and Simulation Approaches: A Systematic Review of Published Models, Applications, and Model Verification," 2015. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4613950/.

[78] WHO, "Harmonization Project Document No. 9; Characterization And Application Of Physiologically Based Pharmacokinetic Models In Risk Assessment," 2010. [Online]. Available: http://www.inchem.org/documents/harmproj/harmproj/harmproj9.pdf.

[79] D. Mukherjee, S. G. Royce, J. A. Alexander, B. Buckley, S. S. Isukapalli, E. V. Bandera, H. Zarbl and P. G. Georgoloulos, "Physiologically-Based Toxicokinetic Modeling of Zearalenone and Its Metabolites: Application to the Jersey Girl Study," 2014.

[80] K. Krishnan and T. Peyret, "Physiologically Based Toxicokinetic (PBTK) Modeling in

Ecotoxicology," 2009. [Online]. Available: https://rd.springer.com/chapter/10.1007/978-1-4419-0197-2_6.

[81] J. DeJongh, M. Nordin-Andersen, B. A. Ploeger and A. Forsby, "Estimation of systemic toxicity of acrylamide by integration of in vitro data with kinetic simulations," *Toxicol Appl Pharmacol,* vol. 158, p. 261–8, 1999.

[82] A. Forsby and B. Blaauboer, "Integrated in vitro approaches for assessing systemic toxicity," in *Proceedings of the symposium Alternative Toxicological Methods for the New Millennium: Science and Application*, Bethesda, 28 November - 1 December 2000, 2001.

[83] G. Eisenbrand, B. Pool-Zobel, V. Baker, M. Balls, B. J. Blaauboer, A. Boobis, A. Carere, S. Kevekordes, J.-C. Lhuguenot, R. Pieters and J. Kleiner, "Methods of In vitro Toxicology," *Food Chem Toxicol,* vol. 40, pp. 193-236, 2002.

[84] A. Forsby and B. Blaauboer, "Integration of in vitro neurotoxicity data with biokinetic modelling for the estimation of in vivo neurotoxicity," *Human Exp Toxicol,* vol. 26, p. 333–8, 2007.

[85] J. Louisse, E. de Jong, J. J. van de Sandt, B. J. Blaauboer, R. A. Woutersen, A. H. Piersma, I. M. Rietjens and M. Verwei, "The use of in vitro toxicity data and physiologically based kinetic modeling to predict dose-response curves for in vivo developmental toxicity of glycol ethers in rat and man.," *Toxicol Sci,* vol. 118, no. 2, pp. 470-84, 2010.

[86] W. M. Gubbels-van Hal, B. J. Blaauboer, H. M. Barentsen, M. A. Hoitink, I. A. Meerts and J. C. van der Hoeven, "An alternative approach for the safety evaluation of new and existing chemicals, an exercise in integrated testing," *Regul Toxicol Pharmacol,* vol. 42, pp. 284-95, 2005.

[87] J. Louisse, S. Bosgra, B. J. Blaauboer, I. M. C. M. Rietjens and M. Verwei, "Prediction of in vivo developmental toxicity of all-trans-retinoic acid based on in vitro toxicity data and in silico physiologically based kinetic modeling," *Arch Toxicol,* vol. 89, no. 7, p. 1135–48, 2015.

[88] J. F. Wambaugh, B. A. Wetmore, R. Pearce, C. Strope, R. Goldsmith, J. P. Sluka, A. Sedykh, A. Tropsha, S. Bosgra, I. Shah, R. Judson, R. S. Thomas and R. W. Setzer, "Toxicokinetic Triage for Environmental Chemicals," *Toxicol Sci,* vol. 147, no. 1, pp. 55-67, 2015.

[89] B. A. Wetmore, J. F. Wambaugh, B. Allen, S. S. Ferguson, M. A. Sochaski, R. W. Setzer, K. A. Houck, C. L. Strope, K. Cantwell, R. S. Judson, E. LeCluyse, H. J. Clewell, R. S. Thomas and M. E. Andersen, "Incorporating High-Throughput Exposure Predictions With Dosimetry-Adjusted In Vitro Bioactivity to Inform Chemical Toxicity Testing," *Toxicol Sci,* vol. 148, no. 1, pp. 121-36, 2015.

[90] L. Turco, T. Catone, F. Caloni, E. Di Consigliio, E. Testai and A. Stammati, "Caco-2/TC7 cell line characterization for intestinal absorption: How reliable is this in vitro model for the prediction of the oral dose fraction absorbed in human?," 2011. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0887233310002122?via%3Dihub.

[91] C. A. Lipinski, F. Lombardo, B. W. Dominy and P. J. Feeney, "Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings," *Adv Drug Deliv Rev,* vol. 23, no. 1-3, pp. 3-25, 1997.

[92] A. K. Ghose, V. N. Viswanadhan and J. J. Wendoloski, "A Knowledge-Based Approach in Designing Combinatorial or Medicinal Chemistry Libraries for Drug Discovery. 1. A Qualitative and Quantitative Characterization of Known Drug Databases," *J Comb Chem,* vol. 1, no. 1, pp. 55-68, 1998.

[93] W. J. Egan, K. M. Merz and J. J. Baldwin, "Prediction of drug absorption using multivariate statistics. Journal of Medicinal Chemistry," *J Med Chem,* vol. 43, no. 21, p. 3867–77, 2000.

[94] T. I. Oprea, "Property distribution of drug-related chemical databases," *J Comput Aided Mol Design,* vol. 14, no. 3, pp. 251-64, 2000.

[95] Y. J. Martin, "A bioavailability score," *J Med Chem,* vol. 48, no. 9, pp. 3164-70, 2005.

[96] T. W. Johnson, K. R. Dress and M. Edwards, "Using the Golden Triangle to optimize clearance and oral absorption," *Bioorg Med Chem Lett,* vol. 19, no. 19, pp. 5560-4, 2009.

[97] European Commission, "Guidance Document on Dermal Absorption (Sanco/222/2000 rev. 7)," 2004. [Online]. Available: https://ec.europa.eu/food/sites/food/files/plant/docs/pesticides_ppp_app-proc_guide_tox_dermal-absorp-2004.pdf.

[98] C. De Heer, A. Wilschut, H. Stevenson and B. C. Hackkert, "Guidance document on the estimation of dermal absorption according to a tiered approach : an update (TNO Report V98.1237)," 1999.

[99] OECD, "Series on Testing and Assessment No. 28, Guidance document for the conduct of skin absorption studies (ENV/JM/MONO(2004)2)," 2004. [Online]. Available: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?doclanguage=en&cote=env/jm/mono%282004%292.

[100] OECD, "OECD Series on Testing and Assessment No. 156, Guidance notes on dermal absorption (ENV/JM/MONO(2011)36)," 2011. [Online]. Available: https://www.oecd.org/chemicalsafety/testing/48532204.pdf.

[101] EFSA, "Guidance on Dermal Absorption," *EFSA Journal ,* vol. 10, no. 4, p. 2665, 2012.

[102] European Commission / SCCS (Scientific Committee on Consumer Safety), "Basic criteria for the in vitro assessment of dermal absorption of cosmetic ingredients (SCCS/1358/10)," 2010. [Online]. Available: https://ec.europa.eu/health/scientific_committees/consumer_safety/docs/sccs_s_002.pdf.

[103] I. Tsakovska, I. Pajeva, M. Al Sharif, P. Alov, E. Fioravanzo, S. Kovarich, A. P. Worth, A. N. Richarz, C. Yang, A. Mostrag-Szlichtyng and M. T. D. Cronin, "Quantitative structure-skin permeability relationships," *Toxicology,* vol. 387, pp. 27-42, 2017.

[104] G. L. Flynn, Principles of route-to-route extrapolation for risk assessment, New York: Elsevier, 1990, pp. 93-127.

[105] B. M. Magnusson, W. J. Pugh and M. S. Roberts, "Simple rules defining the potential of compounds for transdermal delivery or toxicity.," *Pharm Res,* vol. 21, no. 6, pp. 1047-54, 2004.

[106] H. Van de Waterbeemed, G. Camenisch, G. Folkers, J. Chretien and O. A. Raevsky, "Estimation of blood-brain barrier crossing of drugs using molecular size and shape, and H-bonding descriptors," *J Drug Target,* vol. 6, no. 2, pp. 151-65, 1998.

[107] J. C. Madden, "Toxicokinetics Considerations in predicting toxicity," in *In silico Toxicology: Principles and Applications*, M. Cronin and J. Madden, Eds., Cambridge, RSC publishing, 2010, pp. 531-57.

[108] A. M. Lostia, M. C. Zorzoli, S. Fortaner Torrent, E. Mendoza, M. Bouhfid and S. Coecke, "Development of an in vitro metabolic hepatic clearance method (JRC Technical reports EUR 28354 EN)," Publications Office of the European Union, Luxembourg, 2016.

[109] R. M. Rodrigues, J. De Kock, S. Branson, M. Vinken, K. Meganathan, U. Chaudhari, A. Sachinidis, G. O, T. Roskams, V. De Boe, T. Vanhaecke and V. Rogiers, "Human Skin-Derived Stem Cells as a Novel Cell Source for In Vitro Hepatotoxicity Screening of Pharmaceuticals," 2014. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3870603/.

[110] R. M. Rodrigues, H. A. V. De Boe, S. A. U. Chaudhari, O. Govaere, T. Roskams, T. Vanhaecke, V. Rogiers and J. De Kock, "Toxicogenomics-based prediction of acetaminophen-induced liver injury using human hepatic cell systems," 2016. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/26497421.

[111] J. H. Sung and M. Shuler, "In vitro microscale systems for systematic drug toxicity study," *Bioprocess Biosyst Eng,* vol. 33, no. 1, pp. 5-19, 2010.

[112] F. Khalil and S. Läer, "Physiologically Based Pharmacokinetic Modeling: Methodology, Applications, and Limitations with a Focus on Its Role in Pediatric Drug Development," 2011. [Online]. Available: https://www.hindawi.com/journals/bmri/2011/907461/.

[113] S. Coecke, O. Pelkonen, S. Batista Leite, U. Bernauer, J. G. M. Bessems, F. Y. Boi, U. Gundert-Remy, G. Loizou, E. Testai, Zaldivar and J-M, "Toxicokinetics as a key to the integrated toxicity risk assessment based primarily on non-animal approaches," 2013. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0887233312001622?via%3Dihub.

[114] The National Academies of Sciences, Engineering, and Medicine, "Non testing Approaches Relevant to Prediction of Acute Toxicity and Potency," in *Application of Modern Toxicology Approaches for Predicting Acute Toxicity for Chemical Defense*, Washington, DC, The National Academies Press, 2015, pp. 29-50.

[115] A. Kopp-Schneider, P. Prieto, A. Kinsner-Ovaskainen and S. Stanzel, "Design of a testing strategy using non-animal based test methods: Lessons learnt from the ACuteTox project," *Toxicology in Vitro,* vol. 27, no. 4, pp. 1395-401, 2013.

[116] J. Hamm, K. Sullivan, A. J. Clippinger, J. Strickland, S. Bell, B. Bhhatarai, B. Blaauboer, W. Casey, D. Dorman, A. Forsby, N. Garcia-Reyero, S. Gehen, R. Graepel, J. Hotchkiss, A. Lowit, J. Matheson, E. Reaves, L. Scarano, C. Sprankle, J. Tunkel, D. Wilson, M. Xia, H. Zhu and D. Allen, "Alternative approaches for identifying acute systemic toxicity: Moving from research to regulatory testing.," *Toxicol In Vitro,* vol. in press, 2017.

[117] E. Houthoff, P. Rugen and D. Hart, "Predictability of in vitro dermal assays when evaluating fatty amine derivatives," *Toxicol In Vitro,* vol. 29, pp. 1263-7, 2015.

[118] OECD, "Series on Testing and Assessment No. 160, Guidance document on the bovine corneal opacity and permeability (Bcop) and isolated chicken eye (Ice) test methods: Collection of tissues for histological evaluation and collection of data on non-severe irritants," 2011. [Online]. Available: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=ENV/JM/MONO%282011%2945&doclanguage=en.

[119] OECD, "OECD Series on Testing and Assessment, No. 168. The Adverse Outcome Pathway for Skin Sensitisation Initiated by Covalent Binding to Proteins Part 1: Scientific Evidence (ENV/JM/MONO(201)10/PART1)," 2012. [Online]. Available: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono(2012)10/part1&doclanguage=en.

[120] C. A. Redlich, "Skin exposure and asthma: is there a connection," *Proc Am Thorac Soc,* vol. 7, pp. 134-7, 2010.

[121] I. Kimber, R. J. Dearman, D. A. Basketter and D. R. Boverhof, "Chemical respiratory respiratory allergy: reverse engineering an adverse outcome pathway," *Toxicology,* vol. 318, pp. 32-9, 2015.

[122] J. S. Jaworska, A. Natsch, C. Ryan, J. Strickland, T. Ashikaga and M. Miyazawa, "A Bayesian integrated testing strategy (ITS) for skin sensitization potency assessment: a decision support system for quantitative weight of evidence and adaptive testing strategy," *Arch Toxicol,* vol. 89, no. 12, pp. 2355-83, 2015.

[123] A. Natsch, R. Emter, H. Gfeller, T. Haupt and G. Ellis, "Predicting skin sensitizer potency based on in vitro data from KeratinoSens and kinetic peptide binding: global versus domain-based assessment," *Toxicol Sci,* vol. 43, no. 2, pp. 319-32, 2015.

[124] O. Takenouchi, S. Fukui, K. Okamoto, S. Kurotani, N. Imai, M. Fujishiro, D. Kyotani, Y. Kato, T. Kasahara, M. Fujita, A. Toyoda, D. Sekiya, S. Watanabe, H. Seto, M. Hirota, T. Ashikaga and M. Miyazawa, "Test battery with the human cell line activation test, direct peptide reactivity assay and DEREK based on a 139 chemical data set for predicting skin sensitizing potential and potency of chemicals," *J Appl Toxicol,* vol. 35, no. 11, pp. 1318-32, 2015.

[125] OECD, "Annex I: Case Studies To The Guidance Document On The Reporting Of Defined Approaches And Individual Information Sources To Be Used Within Integrated

Approaches To Testing And Assessment (IATA) For Skin Sensitisation," 2016. [Online]. Available: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono%282016%2929/ann1&doclanguage=en.

[126] OECD, "Annex II: Information Sources To Be Used Within The Case Studies To The Guidance Document On The Reporting Of Defined Approaches And Individual Information Sources To Be Used Within IATA For Skin Sensitisation," 2016.

[127] K. Reisinger, S. Hoffmann, N. Alépée, T. Ashikaga, J. Barroso, C. Elcombe, N. Gellatly, V. Galbiati, S. Gibbs, H. Groux, J. Hibatallah, D. Keller, P. Kern, M. Klaric, S. Kolle, J. Kuehnl, N. Lambrechts, M. Lindstedt, M. Millet, S. Martinozzi-Teissier, A. Natsch, D. Petersohn, I. Pike, H. Sakagutchi, A. Schepky, M. Tailhardat, M. Templier, E. van Vliet and G. Maxwell, "Systematic evaluation of non-animal test methods for skin sensitisation safety assessment," *Toxicol in Vitro,* vol. 29, no. 1, pp. 259-70, 2015.

[128] D. Asturiol, S. Casati and A. Worth, "Consensus of classification trees for skin sensitisation hazard prediction," *Toxicol in Vitro,* vol. 36, p. 197–209, 2016.

[129] R. Kroes, A. G. Renwick, M. Cheeseman, J. Kleiner, I. Mangelsdorf, A. Piersma, B. Schilter, J. Schlatter, F. van Schothorst, J. G. Vos and G. Würtzen, "Structure-based thresholds of toxicological concern (TTC): guidance for application to substances present at low levels in the diet," *Food Chem Toxicol,* vol. 42, no. 1, pp. 65-83, 2004.

[130] Temanord, " Threshold of Toxicological Concern (TTC) - Literature review and applicability," 2005. [Online]. Available: http://norden.diva-portal.org/smash/get/diva2:702709/FULLTEXT01.pdf.

[131] S. Lapenna, M. Fuart-Gatnik and A. Worth, "Review of QSAR Models and Software Tools for predicting Acute and Chronic Systemic Toxicity, JRC Scientific and Technical Report (EUR 24639 EN - 2010)," 2010. [Online]. Available: http://publications.jrc.ec.europa.eu/repository/bitstream/JRC61930/eur_24639_en.pdf.

[132] E. J. Matthews, N. L. Kruhlak, J. L. Weaver, R. D. Benz and J. F. Contrera, "Assessment of the health effects of chemicals in humans: II. Construction of an adverse effects database for QSAR modeling," *Curr Drug Discov Technol,* vol. 1, pp. 243-54, 2004.

[133] C. J. Ursem, N. L. Kruhlak, J. F. Contrera, P. M. MacLaughlin, R. D. Benz and E. J. Matthews, "Identification of structure-activity relationships for adverse effects of pharmaceuticals in humans. Part A: use of FDA post-market reports to create a database of hepatobiliary and urinary tract toxicities," *Regul Toxicol Pharmacol,* vol. 54, pp. 1-22, 2009.

[134] E. J. Matthews, C. J. Ursem, N. L. Kruhlak, R. D. Benz, D. A. Sabaté, C. Yang, G. Klopman and J. F. Contrera, "Identification of structure-activity relationships for adverse effects of pharmaceuticals in humans: Part B. Use of (Q)SAR systems for early detection of drug-induced hepatobiliary and urinary tract toxicities," *Regul Toxicol Pharmacol,* vol. 54, pp. 23-42, 2009.

[135] D. Kirkland and G. Speit, "Evaluation of the ability of a battery of three in vitro genotoxicity tests to discriminate rodent carcinogens and non-carcinogens III. Appropriate follow-up testing in vivo," *Mutat Res,* vol. 654, no. 2, pp. 114-32, 2008.

[136] S. Hamada, S. Sutou, T. Morita, A. Wakata, S. Asanami, S. Hosoya, S. Ozawa, K. Kondo, M. Nakajima, H. Shimada, K. Osawa, Y. Kondo, N. Asano, S. Sato, H. Tamura, N. Yajima, R. Marshall, C. Moore, D. Blakely, L. Schectman, J. Weaver, D. Torous, R. Proudlock, S. Ito, C. Namiki and M. Hayashi, "Evaluation of the Rodent Micronucleus Assay by a 28-Day Treatment Protocol: Summary of the 13th Collaborative Study by the Collaborative Study Group for the Micronucleus Test (CSGMT)/Environmental Mutagen Society of Japan (JEMS)-Mammalian Mutagenicity Stu," *Environ Mol Mutagen,* vol. 37, no. 2, pp. 93-110, 2001.

[137] S. Pfuhler, M. Fellows, J. van Benthem, R. Corvi, R. Curren, K. Dearfield, P. Fowler, R. Frötschl, A. Elhajouji, L. Le Hégarat, T. Kasamatsu, H. Kojima, G. Ouédraogo, A. Scott and G. Speit, "In vitro genotoxicity test approaches with better predictivity: summary

of an IWGT workshop," *Mutat Res,* vol. 723, no. 2, pp. 101-7, 2009.

[138] D. E. Bowen, J. H. Whitwell, L. Lillford, D. Henderson, D. Kidd, S. Mc Garry, G. Pearce, C. Beevers and D. J. Kirkland, "Evaluation of a multi-endpoint assay in rats, combining the bone-marrow micronucleus test, the Comet assay and the flow-cytometric peripheral blood micronucleus test," *Mutat Res,* vol. 722, no. 1, pp. 7-19, 2011.

[139] E. Madrigal-Bujaidar, E. O. Madrigal-Santillán, I. Alvarez-Gonzalez, R. Baez and F. Marquez, "Micronuclei induced by imipramine and desipramine in mice: a subchronic study," *Basic Clin Pharmacol Toxicol,* vol. 103, no. 6, pp. 569-73, 2008.

[140] A. Rothfuss, M. Honma, A. Czich, M. J. Aardema, B. Burlinson, S. Galloway, S. Hamada, D. Kirkland, R. H. Heflich, J. Howe, M. Nakajima, M. O'Donovan, U. Plappert-Helbig, C. Priestley, L. Recio, M. Schuler, Y. Uno and H. J. Martus, "Improvement of in vivo genotoxicity assessment: combination of acute tests and integration into standard toxicity testing," *Mutat Res,* vol. 723, no. 2, pp. 108-20, 2011.

[141] EFSA, "Scientific opinion on genotoxicity testing strategies applicable to food and feed safety assessment," *EFSA Journal,* vol. 9, no. 9, p. 2379 [69 pp.], 2011.

[142] C. L. Yauk, M. J. Aardema, B. J. B. van Benthem J, K. L. Dearfield, D. M. DeMarini, Y. E. Dubrova, M. Honma, J. R. Lupski, F. Marchetti, M. L. Meistrich, F. Pacchierotti, J. Stewart, M. D. Waters and G. R. Douglas, "Approaches for identifying germ cell mutagens: Report of the 2013 IWGT workshop on germ cell assays," *Mutat Res Genet Toxicol Environ Mutagen,* vol. 783, pp. 36-54, 2015.

[143] M. Hayashi, J. T. MacGregor, D. G. Gatehouse, I. D. Adler, D. H. Blakey, S. D. Dertinger, G. Krishna, T. Morita, A. Russo and S. Sutou, "In vivo rodent erythrocyte micronucleus assay. II. Some aspects of protocol design including repeated treatments, integration with toxicity testing, and automated scoring," *Environ Mol Mutagen.,* vol. 35, no. 3, pp. 234-52, 2000.

[144] R. Benigni and C. Bossa, "Structure alerts for carcinogenicity, and the Salmonella assay system: a novel insight through the chemical relational databases technology.," *Mutat Res,* vol. 659, no. 3, pp. 248-61, 2008.

[145] H. Ellinger-Ziegelbauer, J. Aubrecht, J. C. Kleinjans and H. J. Ahr, "Application of toxicogenomics to study mechanisms of genotoxicity and carcinogenicity," *Toxicol Lett,* vol. 186, no. 1, pp. 36-44, 2009.

[146] M. D. Waters, M. Jackson and I. Lea, "Characterizing and predicting carcinogenicity and mode of action using conventional and toxicogenomics methods," *Mutat Res,* vol. 705, no. 3, pp. 184-200, 2010.

[147] T. Y. Doktorova, R. Yildirimman, M. Vinken, M. Vilardell, T. Vanhaecke, H. Gmuender, R. Bort, G. Brolen, G. Holmgren, R. Li, C. Chesne, J. van Delft, J. Kleinjans, J. Castell, P. Bjorquist, R. Herwig and V. Rogiers, "Transcriptomic responses generated by hepatocarcinogens in a battery of liver-based in vitro models," *Carcinogenesis,* vol. 34, no. 6, pp. 1393-402, 2013.

[148] P. Jennings, A. Limonciel, L. Felice and M. O. Leonard, "An overview of transcriptional regulation in response to toxicological insult," *Arch Toxicol,* vol. 87, no. 1, pp. 49-72, 2013.

[149] D. Kirkland, M. Aardema, L. Henderson and L. Müller, "Evaluation of the ability of a battery of three in vitro genotoxicity tests to discriminate rodent carcinogens and non-carcinogens I. Sensitivity, specificity and relative predictivity," *Mutat Res,* vol. 584, no. 1-2, pp. 1-256, 2005.

[150] D. Kirkland, S. Pfuhler, D. Tweats, M. Aardema, R. Corvi, F. Darroudi, A. Elhajouji, H. Glatt, P. Hastwell, M. Hayashi, P. Kasper, S. Kirchner, A. Lynch, D. Marzin, D. Maurici, J. R. Meunier, L. Müller, G. Nohynek, J. Parry, E. Parry, V. Thybaud, R. Tice, J. van Benthem, P. Vanparys and P. White, "How to reduce false positive results when undertaking in vitro genotoxicity testing and thus avoid unnecessary follow-up animal tests: Report of an ECVAM Workshop," *Mutat Res,* vol. 628, no. 1, pp. 31-55, 2005; 2007.

[151] R. Corvi, F. Madia, A. Worth and M. Whelan, "EURL ECVAM Strategy to Avoid and Reduce Animal Use in Genotoxicity Testing (JRC Scientific and Policy Report EUR 26375 EN)," Publications Office of the European Union, Luxembourg, 2013.

[152] D. Kirkland, L. Reele, D. Gatehouse and P. Vanparys, "A core in vitro genotoxicity battery comprising the Ames test plus the in vitro micronucleus test is sufficient to detect rodent carcinogens and in vivo genotoxins," *Mutat Res,* vol. 721, no. 1, pp. 27-73, 2011.

[153] EMA, " International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use guideline S2 (R1) on genotoxicity testing-Step 5," 2012. [Online]. Available: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2011/12/WC500119604.pdf.

[154] UK COM, "Guidance on a strategy for genotoxicity testing of chemical substances," 2011. [Online]. Available: https://www.gov.uk/government/organisations/committee-on-mutagenicity-of-chemicals-in-food-consumer-products-and-the-environment.

[155] European Commission / SCCS (Scientific Committee on Consumer Safety), "The SCCS Notes of Guidance for the Testing of Cosmetic Ingredients and their Safety Evaluation, 9th revision, adopted on 29 September 2015 and published on 25 April 2016. SCCS/1564/15.," 2016. [Online]. Available: http://ec.europa.eu/health//sites/health/files/scientific_committees/consumer_safety/docs/sccs_o_190.pdf.

[156] D. Kirkland, E. Zeiger, F. Madia and R. Corvi, "Can in vitro mammalian cell genotoxicity test results be used to complement positive results in the Ames test and help predict carcinogenic or in vivo genotoxic activity? II. Construction and analysis of a consolidated database.," *Mutat Res Genet Toxicol Environ Mutagen,* Vols. 775-776, pp. 69-80, 2014.

[157] T. Hu, Z. S. Khambatta, P. J. Hayden, J. Bolmarcich, R. L. Binder, M. K. Robinson, G. J. Carr, J. P. Tiesman, B. B. Jarrold, R. Osborne, T. D. Reichling, S. T. Nemeth and M. J. Aardema, "Xenobiotic metabolism gene expression in the EpiDermin vitro 3D human epidermis model compared to human skin," *Toxicol In Vitro,* vol. 24, no. 5, pp. 1450-63, 2010.

[158] D. Maurici, M. Aardema, R. Corvi, M. Kleber, C. Krul, C. Laurent, N. Loprieno, M. Pasanen, S. Pfuhler, B. Phillips, E. Sabbioni, T. Sanner and P. Vanparys, "Genotoxicity and mutagenicity," *Altern Lab Anim,* vol. 33, no. Suppl 1, pp. 117-30, 2005.

[159] S. Pfuhler, A. Kirst, M. Aardema, N. Banduhn, C. Goebel, D. Araki, M. Costabel-Farkas, E. Dufour, R. Fautz, J. Harvey, N. J. Hewitt, J. Hibatallah, P. Carmichael, M. Macfarlane, K. Reisinger, J. Rowland, F. Schellauf, A. Schepky and J. Scheel, "A tiered approach to the use of alternatives to animal testing for the safety assessment of cosmetics: genotoxicity. A COLIPA analysis," *Regul Toxicol Pharmacol,* vol. 57, no. 2-3, pp. 315-24, 2010.

[160] M. Kirsch-Volders, I. Decordier, A. Elhajouji, G. Plas, M. J. Aardema and M. Fenech, "In vitro genotoxicity testing using the micronucleus assay in cell lines, human lymphocytes and 3D human skin models," *Mutagenesis,* vol. 26, no. 1, pp. 177-84, 2011.

[161] T. Wolf, C. Niehaus-Rolf, N. Banduhn, D. Eschrich, J. Scheel and N. P. Luepke, "The hen's egg test for micronucleus induction (HET-MN): novel analyses with a series of well-characterized substances support the further evaluation of the test system," *Mutat Res,* vol. 650, no. 2, pp. 150-64, 2008.

[162] R. Corvi and F. Madia, "In vitro genotoxicity testing - Can the performance be enhanced?," *Food Chem Toxicol,* Vols. S0278-6915, no. 16, pp. 30290-3, 2016.

[163] C. Sonich-Mullin, R. Fielder, J. Wiltse, K. Baetcke, J. Dempsey, P. Fenner-Crisp, D. Grant, M. Hartley, A. Knaap, D. Kroese, I. Mangelsdorf, E. Meek, J. M. Rice and M. Younes, "IPCS conceptual framework for evaluating a mode of action for chemical

carcinogenesis," *Regul Toxicol Pharmacol,* vol. 34, no. 2, pp. 146-52, 2001.

[164] A. R. Boobis, S. M. Cohen, V. Dellarco, D. McGregor, M. E. Meek, C. Vickers, D. Willcocks and W. Farland, "IPCS framework for analyzing the relevance of a cancer mode of action for humans," *Crit Rev Toxicol,* vol. 36, no. 10, pp. 781-92, 2006.

[165] WHO, "WHO Harmonization Project Document No. 4 PART 1: IPCS Framework For Analysing The Relevance Of A Cancer Mode Of Action For Humans And Case-Studies, PART 2: IPCS Framework For Analysing The Relevance Of A Non-Cancer Mode Of Action For Humans," 2007. [Online]. Available: http://www.who.int/ipcs/methods/harmonization/areas/cancer_mode.pdf?ua=1.

[166] S. J. Enoch and M. T. Cronin, "A review of the electrophilic reaction chemistry involved in covalent DNA binding," *Crit Rev Toxicol,* vol. 40, no. 8, pp. 728-48, 2010.

[167] G. Patlewicz, N. Jeliazkova, A. Gallegos Saliner and A. P. Worth, "Toxmatch-a new software tool to aid in the development and evaluation of chemically similar groups," *SAR QSAR Environ Res,* vol. 19, no. 3-4, pp. 397-412, 2008.

[168] J. A. DiPaolo, P. Donovan and R. Nelson, "Quantitative studies of in vitro transformation by chemical carcinogens," *J Natl Cancer Inst,* vol. 42, no. 5, pp. 867-74, 1969.

[169] R. J. Isfort and R. A. LeBoeuf, "The Syrian hamster embryo (SHE) cell transformation system: a biologically relevant in vitro model--with carcinogen predicting capabilities--of in vivo multistage neoplastic transformation," *Crit Rev Oncog,* vol. 6, no. 3-6, pp. 251-60, 1995.

[170] E. J. Matthews, J. W. Spalding and R. W. Tennant, "Transformation of BALB/c-3T3 cells: V. Transformation responses of 168 chemicals compared with mutagenicity in Salmonella and carcinogenicity in rodent bioassays," *Environ Health Perspect,* vol. 101, no. Suppl 2, pp. 347-482, 1993.

[171] R. A. LeBoeuf, K. A. Kerckaert, M. J. Aardema and R. J. Isfort, "Use of Syrian hamster embryo and BALB/c 3T3 cell transformation for assessing the carcinogenic potential of chemicals," *IARC Sci Publ,* vol. 146, pp. 409-25, 1999.

[172] OECD, "Series on Testing and Assessment No. 31, Detailed review paper on cell transformation assays for detection of chemical carcinogens (ENV/JM/MONO(2007)18)," 2007. [Online]. Available: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono%282007%2918&doclanguage=en.

[173] OECD, "Series on Testing and Assessment No. 214, Guidance Document on the In Vitro Syrian Hamster Embryo (SHE) Cell Transformation Assay (ENV/JM/MONO(2015)18)," 2015. [Online]. Available: http://www.oecd.org/env/ehs/testing/Guidance-Document-on-the-in-vitro-Syrian-Hamster-Embryo-Cell-Transformation-Assay.pdf.

[174] OECD, "Series on Testing and Assessment No. 231, Guidance Document on The In Vitro Bhas 42 Cell Transformation Asssay (ENV/JM/MONO(2016)1)," 2016. [Online]. Available: https://www.oecd.org/env/ehs/testing/ENV_JM_MONO%282016%291.pdf.

[175] E. Benfenati, R. Benigni, D. M. Demarini, C. Helma, D. Kirkland, T. M. Martin, P. Mazzatorta, G. Ouédraogo-Arras, A. M. Richard, B. Schilter, W. G. Schoonen, R. D. Snyder and C. Yang, "Predictive models for carcinogenicity and mutagenicity: frameworks, state-of-the-art, and perspectives," *J Environ Sci Health C Environ Carcinog Ecotoxicol Rev,* vol. 27, no. 2, pp. 57-90, 2009.

[176] A. A. Toropov, A. P. Toropova and E. Benfenati, "Additive SMILES-based carcinogenicity models: Probabilistic principles in the search for robust predictions," *Int J Mol Sci,* vol. 10, no. 7, pp. 3106-27, 2009.

[177] CCCF (Committee on Carcinogenicity of Chemicals in Food, Consumer Products and the Environment), "Guidance on a Strategy for the Risk Assessment of Chemical Carcinogens," 2004. [Online]. Available: https://cot.food.gov.uk/sites/default/files/cot/cocsection.pdf.

[178] OECD, "OECD Series on Testing and Assessment No. 35 and Series on Pesticides No. 14, Guidance notes for analysis and evaluation of chronic toxicity and carcinogenicity studies (ENV/JM/MONO(2002)19)," 2002. [Online]. Available: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono%282002%2919&doclanguage=en.

[179] D. Jacobson-Kram, F. D. Sistare and A. Jacobs, "Use of transgenic mice in carcinogenicity hazard assessment," *Toxicol Pathol,* vol. 32, no. Suppl 1, pp. 49-52, 2004.

[180] J. B. Pritchard, J. E. French, B. J. Davis and J. K. Haseman, "The role of transgenic mouse models in carcinogen identification," *Environ Health Perspect,* vol. 111, no. 4, pp. 444-54, 2003.

[181] ILSI/HESI , "Alternatives to carcinogenicity testing project," *Toxicol Pathol ,* vol. 29, no. Suppl, pp. 1-351, 2001.

[182] OECD, "OECD Series on Testing and Assessment No. 116, Guidance Document on the Conduct and Design of Chronic Toxicity and Carcinogenicity Studies, Supporting Test Guidelines 451, 452 and 453 (ENV/JM/MONO(2011)47)," 2011. [Online]. Available: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=ENV/JM/MONO%282011%2947&doclanguage=en.

[183] M. Widschwendter and P. A. Jones, "DNA methylation and breast carcinogenesis," *Oncogene,* vol. 21, no. 35, pp. 5462-82, 2002.

[184] S. B. Baylin and J. E. Ohm, "Epigenetic gene silencing in cancer - a mechanism for early oncogenic pathway addiction?," *Nat Rev Cancer,* vol. 6, no. 2, pp. 107-16, 2006.

[185] M. Esteller, "Cancer epigenomics: DNA methylomes and histone-modification maps," *Nat Rev Genet,* vol. 8, no. 4, pp. 286-98, 2007.

[186] H. Ellinger-Ziegelbauer, H. Gmuender, A. Bandenburg and H. J. Ahr, "Prediction of a carcinogenic potential of rat hepatocarcinogens using toxicogenomics analysis of short-term in vivo studies," *Mutat Res,* vol. 637, no. 1-2, pp. 23-39, 2008.

[187] M. R. Fielden, R. Brennan and J. Gollub, "A gene expression biomarker provides early prediction and mechanistic assessment of hepatic tumor induction by nongenotoxic chemicals," *Toxicol Sci,* vol. 99, no. 1, pp. 90-100, 2007.

[188] K. Stemmer, H. Ellinger-Ziegelbauer, H. J. Ahr and D. R. Dietrich, "Carcinogen-specific gene expression profiles in short-term treated Eker and wild-type rats indicative of pathways involved in renal tumorigenesis," *Cancer Res,* vol. 67, no. 9, pp. 4052-68, 2007.

[189] P. Nioi, I. D. Pardo, P. J. Sherratt and R. D. Snyder, "Prediction of non-genotoxic carcinogenesis in rats using changes in gene expression following acute dosing," *Chem Biol Interact,* vol. 172, no. 3, pp. 206-15, 2008.

[190] T. Uehara, M. Hirode, A. Ono, N. Kiyosawa, K. Omura, T. Shimizu, Y. Mizukawa, T. Miyagishima, T. Nagao and T. Urushidani, "A toxicogenomics approach for early assessment of potential non-genotoxic hepatocarcinogenicity of chemicals in rats," *Toxicology,* vol. 250, no. 1, pp. 15-26, 2008.

[191] M. J. Jonker, O. Bruning, M. van Iterson, M. M. Schaap, T. V. van der Hoeven, H. Vrieling, R. B. Beems, A. de Vries, H. van Steeg, T. M. Breit and M. Luijten, "Finding transcriptomics biomarkers for in vivo identification of (non-)genotoxic carcinogens using wild-type and Xpa/p53 mutant mouse models," *Carcinogenesis,* vol. 30, no. 10, pp. 1805-12, 2009.

[192] A. Rohrbeck, G. Salinas, K. Maaser, J. Linge, S. Salovaara, R. Corvi and J. Borlak, "Toxicogenomics applied to in vitro carcinogenicity testing with Balb/c 3T3 cells revealed a gene signature predictive of chemical carcinogens," *Toxicol Sci,* vol. 118, no. 1, pp. 31-41, 2010.

[193] L. Ao, J. Y. Liu, W. B. Liu, L. H. Gao, R. Hu, Z. J. Fang, Z. X. Zhen, M. H. Huang, M. S. Yang and J. Cao, "Comparison of gene expression profiles in BALB/c 3T3 transformed

foci exposed to tumor promoting agents," *Toxicol In Vitro,* vol. 24, no. 2, pp. 430-8, 2010.

[194] M. Vaccari, M. Mascolo, F. Rotondo, E. Morandi, D. Quercioli, S. Perdichizzi, C. Zanzi, S. Serra, V. Poluzzi, P. Angelini, S. Grilli and A. Colacci, "Identification of pathway-based toxicity in the BALB/c 3T3 cell model," *Toxicol In Vitro,* vol. 29, no. 6, pp. 1240-53, 2015.

[195] J. E. Klaunig, L. M. Kamendulis and B. A. Hocevar, "Oxidative stress and oxidative damage in carcinogenesis," *Toxicol Pathol,* vol. 38, no. 1, pp. 96-109, 2010.

[196] J. E. Klaunig and Y. Shi, "UNIT 2.17 Assessment of Gap Junctional Intercellular Communication," *Curr Protoc Toxicol,* vol. 41, pp. 2.17.1-2.17.10., 2009.

[197] M. N. Jacobs, A. M. Colacci, K. Louekari, M. Luijten, B. C. Hakkert, M. Paparella and P. Vasseur, "International regulatory needs for development of an IATA for non-genotoxic carcinogenic chemical substances," 2016. [Online]. Available: http://www.altex.ch/resources/Jacobs_of_v2.pdf.

[198] S. Xiao, J. Coppeta, H. Rogers, B. Isenberg, J. Zhu, S. Olalekan, K. McKinnon, D. Dokic, A. Rashedi, D. Haisenleder, S. Malpani, C. Arnold-Murray, K. Chen, M. Jiang, L. Bai, C. Nguyen, J. Zhang, M. Laronda, T. Hope, K. Maniar, M. Pavone, M. Avram , E. Sefton, S. Getsios, J. Burdette, J. Kim, J. Borenstein and T. Woodruff, "A microfluidic culture model of the human reproductive tract and 28-day menstrual cycle," *Nat Commun,* vol. 8, p. 14584, 2017.

[199] M. Bjorling-Poulsen, H. Raun Andersen and P. Grandjean, "Potential developmental neurotoxicity of pesticides used in Europe," *Environ Health,* vol. 7, no. 50, pp. 1-22, 2008.

[200] P. Grandjean and P. J. Landrigan, "Neurobehavioural effects of developmental toxicity," *The Lancet,* vol. 13, no. 3, pp. 330-338, 2014.

[201] E. Fritsche, H. Alm, J. Baumann, L. Geerts, H. Häkansson, S. Masjosthusmann and H. Witters, "Literature review on in vitro and alternative Developmental Neurotoxicity (DNT) testing methods (EFSA Supporting Publications 2015.EN-778)," 2015. [Online]. Available: http://onlinelibrary.wiley.com/doi/10.2903/sp.efsa.2015.EN-778/abstract.

[202] OECD, "OECD Series on Testing and Assessment No. 261, Report of the OECD/EFSA workshop on developmental neurotoxicity (DNT): the use of non-animal test methods for regulatory purposes ENV/JM/MONO(2017)4," 2017. [Online]. Available: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=ENV/JM/MONO%282017%294&docLanguage=En.

[203] OECD, "Series on Testing and Assessment No. 261: Annex 1 Report of the OECD/EFSA workshop on developmental neurotoxicity (DNT): The use of non-animal test methods for regulatory purposes, Background document (ENV/JM/MONO(2017)4/ANN1)," [Online]. Available: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=ENV/JM/MONO%282017%294/ANN1&docLanguage=En.

[204] A. Beronius, A. Hanberg, R. Heimeier and H. Håkansson, "IMM-rapport 1/2013 Risk assessment of developmental neurotoxicity: Evaluation of the OECD TG 426 test guideline and guidance documents," 2013. [Online]. Available: http://ki.se/sites/default/files/2013-1.pdf.

[205] K. M. Crofton, W. R. Mundy, P. J. Lein, A. Bal-Price, S. Coecke, A. Seiler, H. Knaut, L. Buzanska and A. Goldberg, "Developmental neurotoxicity testing: recommendations for developing alternative methods for the screening and prioritization of chemicals," *Altex,* vol. 28, no. 1, pp. 9-15, 2011.

[206] OECD, "Series on Testing and Assessment No.261: Annex 2 Report of the OECD/EFSA workshop on developmental neurotoxicity (DNT): The use of non-animal test methods for regulatory purposes: Slides of Speakers' Plenary Presentations (ENV/JM/MONO(2017)4/ANN2)," 2017. [Online]. Available: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=ENV/JM/MON

O(2017)4/ANN2&docLanguage=En.

[207] WHO/IPCS, "Harmonization Project Document No. 10, Guidance for immunotoxicity risk assessment for chemicals," 2012. [Online]. Available: http://www.who.int/ipcs/methods/harmonization/areas/guidance_immunotoxicity.pdf?ua=1.

[208] ICH, "ICH Topic S 8 Immunotoxicity Studies for Human Pharmaceuticals. Note for Guidance on Immunotoxicity studies for human pharmaceuticals (CHMP/167235/2004)," 2006. [Online]. Available: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500002851.pdf.

[209] R. R. Dietert, R. A. Etzel, D. Chen, M. Halonen, S. D. Holladay, A. M. Jarabek, K. Landreth, D. B. Peden, K. Pinkerton, R. J. Smialowicz and T. Zoetis, "Workshop to identify critical windows of exposure for children's health: immune and respiratory systems work group summary," *Environ Health Perspect,* vol. 108, no. Suppl 3, pp. 483-90, 2000.

[210] S. D. Holladay and R. J. Smialowicz, "Development of the murine and human immune system: differential effects of immunotoxicants depend on time of exposure.," *Environ Health Perspect,* vol. 108, no. Suppl 3, pp. 463-73, 2000.

[211] H. Van Loveren and A. Piersma, "Immunotoxicological consequences of perinatal chemical exposures," *Toxicol Lett,* vol. 149, no. 1-3, pp. 141-5, 2004.

[212] L. A. Burns-Naas, K. L. Hastings, G. S. Ladics, S. L. Makris, G. A. Parker and M. P. Holsapple, "What's so special about the developing immune system?," *Int J Toxicol,* vol. 27, no. 2, pp. 223-54, 2008.

[213] T. Hartung and E. Corsini, "Immunotoxicology: Challenges in the 21st Century and In Vitro Opportunities," [Online]. Available: http://www.altex.ch/resources/altex_2013_4_411_426_FFTHartungandCorsini1.pdf.

[214] WHO/IPCS, "Global assessment of the state-of-the-science of endocrine disruptors (WHO/PCS/EDC/02.2)," 2002. [Online]. Available: http://www.who.int/ipcs/publications/new_issues/endocrine_disruptors/en/.

[215] OECD, "OECD Series on Testing and Assessment No. 150, Guidance document on standardised Test Guidelines for evaluating chemicals for endocrine disruption ENV/JM/MONO (2012)22," 2012. [Online]. Available: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono%282012%2922&doclanguage=en.

[216] WHO/UNEP, "State of the science of endocrine disrupting chemicals - 2012: An assessment of the state of the science of endocrine disruptors prepared by a group of experts for the United Nations Environment Programme (UNEP) and WHO," 2013. [Online]. Available: http://www.who.int/ceh/publications/endocrine/en/.

[217] B. Ayala-García, M. López-Santibáñez Guevara, M.-C. LI, A. Fuentes-Farías, E. Meléndez-Herrera and G. Gutiérrez-Ospina, "Speciation, Phenotypic Variation and Plasticity: What Can Endocrine Disruptors Tell Us?," *Int J Endocrinol,* vol. 2013, p. 862739, 2013.

[218] European Commission, "COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT AND THE COUNCIL on endocrine disruptors COM(2016)350 final," 2016. [Online]. Available: http://ec.europa.eu/health/endocrine_disruptors/docs/com_2016_350_en.pdf.

[219] European Commission, "Commission Delegated Regulation (EU) …/... of 4.9.2017 setting out scientific criteria for the determination of endocrine-disrupting properties pursuant to Regulation (EU) No 528/2012 (C(2017) 5467 final)," 2017. [Online]. Available: http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=pi_com:C(2017)5467.

[220] European Commission, "Draft Commission Regulation (EU) …/… of XXX setting out scientific criteria for the determination of endocrine disrupting properties and amending Annex II to Regulation (EC) 1107/2009 C(2016)3451 projet," 2016. [Online].

Available: http://ec.europa.eu/health//sites/health/files/endocrine_disruptors/docs/2016_pppcriteria_en.pdf.

[221] European Commission, "Commission Staff Working Document Impact Assessment Defining criteria for identifying endocrine disruptors in the context of the implementation of the plant protection products regulation and biocidal products regulation, Main report SWD(2016) 21 final," 2016. [Online]. Available: http://ec.europa.eu/health//sites/health/files/endocrine_disruptors/docs/2016_impact_assessment_en.pdf.

[222] ASTM, "Standard Guide for Conducting Bioconcentration Tests with Fishes and Saltwater Bivalve Mollusks (E1022 – 94)," 2003. [Online]. Available: ftp://185.72.26.245/Astm/2/01/Section%2011/ASTM1106/PDF/E1022.pdf.

[223] E. A. J. Bleeker and E. M. J. Verbruggen, "Bioaccumulation of polycyclic aromatic hydrocarbons in aquatic organisms (RIVM report 601779002/2009)," 2009. [Online]. Available: Bleeker EAJ and Verbruggen EMJ (2009) Bioaccumulation of polycyclic aromatic hydrocarbons in aquatic organisms. RIVM report 601779002/2009. National Institute for Public Health and 25 the Environment, Bilthoven, The Netherlands.

[224] M. Pavan, A. P. Worth and T. I. Netzeva, "Review of QSAR Models for Bioconcentration," 2006. [Online]. Available: https://eurl-ecvam.jrc.ec.europa.eu/laboratories-research/predictive_toxicology/doc/QSAR_Review_Bioconcentration.pdf.

[225] F. Grisoni, V. Consonni, S. Villa, M. Vighi and R. Todeschini, "QSAR models for bioconcentration: Is the increase in the complexity justified by more accurate predictions?," vol. 127, pp. 171-179, 2015.

[226] G. Treu, W. Drost, U. Jöhncke, C. Rauert and C. Schlechtriem, "The Dessau workshop on bioaccumulation: state of the art, challenges and regulatory implications," *Environ Sci Eur,* vol. 27, no. 1, p. 34, 2015.

[227] Z. Dang, "Interpretation of fish biomarker data for identification, classification, risk assessment and testing of endocrine disrupting chemicals.," *Environ Int,* Vols. 92-93, pp. 422-41, 2016.

[228] OECD, "OECD Series on Testing and Assessment No. 126, Short Guidance On The Threshold Approach For Acute Fish Toxicity (ENV/JM/MONO(2010)17)," 2010. [Online]. Available: https://ntp.niehs.nih.gov/iccvam/suppdocs/feddocs/oecd/oecd-gd126.pdf.

[229] C. R. Elcombe, J. Odum, J. R. Foster, S. Stone, S. Hasmall, A. R. Soames, I. Kimber and J. Ashby, "Prediction of rodent nongenotoxic carcinogenesis: evaluation of biochemical and tissue changes in rodents following exposure to nine nongenotoxic NTP carcinogens," *Environ Health Perspect,* vol. 110, no. 4, pp. 363-75, 2002.

[230] R. Billington, R. W. Lewis, J. M. Mehta and I. Dewhurst, "The mouse carcinogenicity study is no longer a scientifically justifiable core data requirement for the safety assessment of pesticides," *Crit Rev Toxicol,* vol. 40, no. 1, pp. 35-49, 2010.

[231] WHO/IPCS, "WHO/IPCS Environmental Health Criteria (EHC) 180, Principles and Methods for Assessing Direct Immunotoxicity Associated with Exposure to Chemicals," 1996. [Online]. Available: http://www.inchem.org/documents/ehc/ehc/ehc180.htm.

[232] WHO/IPCS, "WHO/IPCS Environmental Health Criteria (EHC) 212, Principles and Methods for Assessing Allergic Hypersensitization Associated with Exposure to Chemicals," 1999. [Online]. Available: http://www.inchem.org/documents/ehc/ehc/ehc212.htm.

[233] WHO/IPCS , "WHO/IPCS Environmental Health Criteria (EHC) 236, Principles and Methods for Assessing Autoimmunity Associated with Exposure to Chemicals," 2007. [Online]. Available: http://www.inchem.org/documents/ehc/ehc/ehc236.pdf.

[234] ECHA, "SVHC Roadmap to 2020," [Online]. Available: https://echa.europa.eu/addressing-chemicals-of-concern/substances-of-potential-concern/svhc-roadmap-to-2020-implementation.

[235] ECHA, "ECHA Screening," [Online]. Available: https://echa.europa.eu/addressing-chemicals-of-concern/substances-of-potential-concern/screening.

[236] European Commission / Joint Research Centre, "Screening methodology to identify potential endocrine disruptors according to different options in the context of an impact assessment, JRC Science For Policy Report (EUR 27955 EN)," 2016. [Online]. Available: http://publications.jrc.ec.europa.eu/repository/bitstream/JRC101950/jrc%20screening%20methodology%20for%20ed%20impact%20assessment%20%28online%29.pdf.

[237] K. P. Friedman, E. D. Watt, M. W. Horning, J. M. Hedge, R. S. Judson, K. M. Crofton, K. A. Houck and S. O. Simmons, "Tiered high-throughout screening approach to identify thyroperoxidase inhibitors within the ToxCast Phase I and II chemical libraries," *Toxicological Sciences,* vol. 151, no. 1, pp. 160-180, 2016.

[238] OECD, "Guidance Document on Aspects of OECD TG 305 on Fish Bioaccumulation, Series on Testing & Assessment No. 264 (ENV/JM/MONO(2017)16)," 2017. [Online]. Available: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=ENV/JM/MONO%282017%2916&doclanguage=en.

[239] OECD, "OECD Series on Testing and Assessment No. 23, Guidance Document on Aquatic Toxicity Testing of Difficult Substances and Mixtures (ENV/JM/MONO(2000)6)," 2000. [Online]. Available: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?doclanguage=en&cote=env/jm/mono%282000%296.

[240] OECD, "OECD Series on Testing and Assessment No. 95, Detailed Review Paper on Fish Life-cycle tests (ENV/JM/MONO(2008)22)," 2008. [Online]. Available: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono%282008%2922&doclanguage=en.

[241] OECD, "OECD Series on Testing and Assessment No. 171, Fish Toxicity Testing Framework (ENV/JM/MONO(2012)16)," 2012. [Online].

APPENDIX 1 List of organisations consulted on the report

The following organisations were consulted on this document before final publication[5]:

European Commission

- DG ENV — Directorate-General for Environment.
- DG GROW — Directorate-General for Internal Market, Industry, Entrepreneurship and SMEs.
- DG SANTE — Directorate-General for Health and Food Safety.
- DG JRC — Directorate-General Joint Research Centre: European Union Reference Laboratory for alternatives to animal testing (EURL ECVAM) Directorate F – Health, Consumers and Reference Materials, Chemical Safety and Alternative Methods Unit (F.3).
- Scientific Committees: Scientific Committee on Consumer Safety (SCCS) and Scientific Committee on Health, Environmental and Emerging Risks (SCHEER).

European Agencies

- EMA — European Medicines Agency.
- EFSA — European Food Safety Authority.

Organisation for Economic Co-operation and Development (OECD)

ECHA Committees

- MSC — Member State Committee.
- RAC — Committee for Risk Assessment.
- BPC — Biocidal Products Committee.

ECHA's Accredited Stakeholder organisations (see full list here)

[5] The comments received from these organisations during the drafting of this document could not be all implemented and that consultation does not necessarily imply full endorsement of the content of this report by these organisations.

APPENDIX 2 List of relevant legislation

Biocidal Products Regulation	Regulation (EU) 528/2012 of the European Parliament and of the Council concerning the making available on the market and use of biocidal products. (see https://echa.europa.eu/regulations/biocidal-products-regulation/legislation)
CLP Regulation	Regulation (EC) No 1272/2008 on classification, labelling and packaging of substances and mixtures. (see https://echa.europa.eu/regulations/clp/legislation)
Protection of Animals Directive	Directive 2010/63/EU of the European Parliament and of the Council of 22 September 2010 on the protection of animals used for scientific purposes and repealing Council Directive of 24 November 1986 on the approximation of laws, regulations and administrative provisions of the Member States regarding the protection of animals used for experimental and other scientific purposes (86/609/EEC). (see http://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1488879009816&uri=CELEX:32010L0063)
REACH Regulation	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. (see https://echa.europa.eu/regulations/reach/legislation)
Test Methods Regulation	Commission Regulation (EC) No 440/2008 of 30 May 2008 laying down test methods pursuant to Regulation (EC) No 1907/2006 of the European Parliament and of the Council on the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH). (see http://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1488884240932&uri=CELEX:02008R0440-20160304)

APPENDIX 3 Information requirements, CLP criteria, relevant test methods and specific adaptation rules per endpoint

1. Toxicokinetics

1.1 Information requirements

1.1.1 Under the REACH Regulation

Toxicokinetics studies *in vivo* are not required under the REACH Regulation, but all available information should be provided, including information from prediction methods.

Annex I, Section 1.0.2 states that "*the human health hazard assessment shall consider the toxicokinetic profile (i.e. absorption, metabolism, distribution and elimination) of the substance*".

Furthermore, REACH Annex VIII states (Section 8.8.1) that an "*assessment of the toxicokinetic behaviour of the substance to the extent that can be derived from the relevant available information*" should be performed.

Even though toxicokinetics is not a toxicological property as such and information is not specifically required to be generated for REACH purposes, the generation of toxicokinetic information can be encouraged as a means to interpret data, assist testing strategy and study design, as well as category development, thus helping to optimise testing.

More detailed information on data requirements under the REACH Regulation can be found in the Guidance on IR&CSA – Chapter R.7c.

1.1.2 For biocide active substances under the BPR

Toxicokinetic data are a core information requirement for biocide active substance approval. Toxicokinetics can be based on existing information (non-human data: physicochemical properties, grouping, (Q)SARs and expert systems, *in vitro* data, human data and/or animal data). It is preferred to generate toxicokinetic data within the required toxicity studies such as repeated-dose toxicity where possible. The standard *in vivo* test battery for toxicokinetics, OECD TG 417 provides the test method for the conduct of toxicokinetic studies either as a standalone test or combined with repeated-dose toxicity studies.

In the absence of *in vivo* data, some of the toxicokinetic data may be derived from *in vitro* experiments. Thus, production of *in vivo* data is not a mandatory requirement under the BPR.

1.2 CLP criteria

Toxicokinetics is not a hazard class under the CLP Regulation.

1.3 Relevant test methods

The following *in vivo*, *in vitro* and *in silico* test methods can be used to investigate toxicokinetics:

In vivo

- Toxicokinetics (OECD TG 417).
- Skin Absorption: *In Vivo* Method (EU B.44/OECD TG 427).

In vitro

- Skin Absorption: *In Vitro* Method (EU B.45/OECD TG 428)

There are various prediction tools to predict the toxicokinetic properties (see for example [10, 75, 88]), but the usefulness of individual toxicokinetic methods can only be judged in the context of the intended application [10]. See also the most commonly used *in vitro* systems for ADME in table 12.2 in [10].

At OECD level, development of a "*Performance-Based Test Guideline for the Establishment on Human-derived hepatic system to investigate biotransformation and toxicity of substances by evaluation of CYP450 induction competence*" has started (project 4.76 [22]).

2. Acute toxicity

2.1 Information requirements

The information needs under the REACH Regulation and the BPR are usually met with internationally accepted test methods, which are acceptable for hazard classification.

For acute toxicity, the hazard classification is based on dose/concentration ranges, where lethality (LD_{50}/LC_{50} values or acute toxicity estimates (ATE)) and specific organ toxicity have been observed (STOT SE).

Concerning "*Specific Target organ toxicity – Single exposure*", according to the CLP Regulation EC 1272/2008, *"[t]he standard animal studies in rats or mice that provide this information are acute toxicity studies, which can include clinical observations and detailed macroscopic and microscopic examination to enable the toxic effects on target tissues/organs to be identified*".

The classification for acute toxicity and specific organ toxicity is usually a result of a WoE approach taking all the information into account and comparing against the classification criteria.

2.1.1 Under the REACH Regulation

Information requirements for acute toxicity depend on the tonnage level and are required for substances manufactured or imported in quantities of one tonne or more per year. Information on one route (oral) is required at Annex VII, Section 8.5.1.

Further information on at least one other route of exposure (inhalation or dermal) is required at higher tonnage levels (to meet the requirements at Annexes VIII-X), depending on the nature of the substance and the likely route of human exposure [Sections 8.5.2 and 8.5.3 and column 2 of section 8.5 of Annex VIII].

A more detailed testing strategy is described in Section R.7.4 of the Guidance on IR&CSA – Chapter R.7a.

2.1.2 For biocide active substances under the BPR

The information needs and testing strategy for acute toxicity are similar to those mentioned above under the REACH Regulation, except that it is not dependent on the amount of substance manufactured or imported.

In addition to the oral route of administration (section 8.7.1 of BPR Annex II), for substances other than gases, the information mentioned under sections 8.7.2 to 8.7.3 of BPR Annex II must be provided for at least one other route of administration.

The choice for the second route will depend on the nature of the substance and the likely route of human exposure. Gases and volatile liquids should be administered by the inhalation route. Dermal toxicity must be reported for an active substance except for gases.

If further testing is needed to assess the potential for acute toxicity by the dermal route, the OECD/EU test methods should be used. In addition, new OECD validated tests for acute dermal toxicity should be taken into account once available in deciding the test strategy.

If the only route of exposure is the oral route, then information for only that route need be provided. If either the dermal or inhalation route is the only route of exposure to humans then an oral test may be considered.

Before a new dermal acute toxicity study is carried out, an *in vitro* dermal penetration study (OECD TG 428) should be conducted to assess the likely magnitude and rate of dermal bioavailability.

There may be exceptional circumstances where data from all routes of administration are deemed necessary. For more information, see Section 8.7 of the Guidance on BPR Volume III - Part A.

2.2 CLP criteria

Acute toxicity hazard categories and acute toxicity estimates (ATEs) defining the respective categories are based on animal data. Categories for specific target organ toxicity – single exposure are based on evidence from humans and/or from experimental animals and dependent on nature and severity of effects and effective dose levels.

2.3 Relevant test methods

Oral route, reduction methods

- Acute Oral Toxicity, Acute Toxic Class Method (EU B.1 tris / OECD TG 423).
- Acute Oral Toxicity – Up-and-Down Procedure (OECD TG 425).
- Acute Oral Toxicity – Fixed Dose Procedure.
- (EU B.1 bis / OECD TG 420).

Tests performed according to an old and deleted EU B.1 / OECD TG 401 are only acceptable under the BPR, if performed before December 2002. According to Section R.7.4 of the Guidance on IR&CSA – Chapter R.7a on acute toxicity, **existing** EU B.1 / OECD TG 401 data would normally be acceptable.

The EU B.1 bis / OECD TG 420 should be considered as the first choice for testing regarding acute toxicity if information requirements are not fulfilled by available information. The information in this test method is 'evident toxicity', clear signs of toxicity that if animals were exposed to a higher concentration they are likely to experience death or severe toxicity.

Inhalation route

- Acute Inhalation Toxicity (EU B.2/OECD TG 403).
- Acute Inhalation Toxicity - Fixed Concentration Procedure (OECD TG 433).
- Acute Inhalation Toxicity – Acute Toxic Class Method (EU B.52/OECD TG 436).

Dermal route

- Acute dermal Toxicity (EU B.3/OECD TG 402).

At OECD level, the updated versions of OECD TG 402 "*Acute dermal toxicity*" and OECD TG 433 "*Acute Inhalation Toxicity – Fixed Concentration Procedure*" were published in October 2017.

2.4 Specific adaptation rules

Under the REACH Regulation and the BPR, *in vivo* testing for acute oral toxicity does not need to be conducted if the substance is classified as **corrosive to the skin**, or if a study on acute toxicity by the inhalation route (as the only route of human exposure) is available, for example, for a gas or a highly volatile substance.

For substances other than gases, information on acute toxicity may also be needed through *a* dermal or inhalation route at Annex VIII level (Section 8.5, column 2 of REACH Annex VIII, and Section R.7.4 of the Guidance on IR&CSA) and for biocide active substances (Section 8.7, column 1 of BPR Annex II, and Section 8.7 of the Guidance on BPR Volume III - Part A).

In addition, according to the revised Section 8.5.3, column 2 of REACH Annex VIII:

"*Testing by the dermal route does not need to be conducted if:*
- *the substance does not meet the criteria for classification as acute toxicity or STOT SE by the oral route and*
- *no systemic effects have been observed in in vivo studies with dermal exposure (e.g. skin irritation, skin sensitisation) or, in the absence of an in vivo study by the oral route, no systemic effects after dermal exposure are predicted on the basis of non-testing approaches (e.g. read across, QSAR studies).*"

3. Skin corrosion/irritation and serious eye damage/eye irritation

3.1 Information requirements

3.1.1 Under the REACH Regulation

For a detailed description of the information requirements for these endpoints under the REACH Regulation, please see Section R.7.2 of the Guidance on IR&CSA – Chapter R.7a.

For skin corrosion/irritation
At REACH Annex VII level, the basic information requirements for skin corrosion and irritation are based on *in vitro* testing only and include an *in vitro* study for skin corrosion (Point 8.1.1) and an *in vitro* study for skin irritation (Point 8.1.2).

At REACH Annex VIII level, only if the *in vitro* studies under Points 8.1.1 and 8.1.2 of Annex VII are not applicable or their results are not adequate for classification and risk assessment, must *in vivo* testing be considered.

For serious eye damage/eye irritation
At REACH Annex VII level, the basic information requirement for serious eye damage/eye

irritation is an *in vitro* study (Point 8.2.1). A second *in vitro* study must be considered if the results from the first *in vitro* study do not allow a conclusive decision on classification for serious eye damage/eye irritation.

At REACH Annex VIII level, only if the *in vitro* studies under Sections 8.2.1 of Annex VII are not applicable or their results not adequate for classification and risk assessment, must *in vivo* testing be considered.

3.1.2 For biocide active substances under the BPR

Under the BPR, the assessment of skin corrosion/irritation and the assessment of serious eye damage/eye irritation are both part of the core dataset (see also Sections 8.1 and 8.2 of the Guidance on BPR Volume III - Part A).

For skin corrosion/irritation
The assessment must follow the sequential testing strategy for dermal irritation and corrosion described in the Appendix to EU Test Guideline B.4 "Acute Toxicity-Dermal Irritation/Corrosion" (Annex B.4 to the Test Methods Regulation (EC) No 440/2008).

For serious eye damage/eye irritation
The assessment must follow the sequential testing strategy for eye irritation and corrosion described in the Appendix to EU Test Guideline B.5 "Acute Toxicity: Eye Irritation/Corrosion" (Annex B.5 to the Test Methods Regulation (EC) No 440/2008).

3.2 CLP criteria

The criteria for the skin corrosive category and sub-categories and the skin irritation category are based on animal data.

In vitro alternatives that have been validated and accepted may also be used to support classification decisions.

For serious eye damage/eye irritation, the classification system involves a tiered testing and evaluation scheme. The criteria themselves for irreversible eye effects or for reversible eye effects are based on animal data.

The CLP Regulation defines in more detail and further specifies duration criteria for the application of the test substance and the observation period to distinguish between skin irritants, which cause reversible damage to the skin following the application of a substance and skin corrosives, which cause irreversible damage to the skin after the same application time and after a two-week observation period (Section 3.2.1.1 of Annex I to the CLP Regulation).

Eye irritation and serious eye damage are defined in the CLP Regulation as fully reversible or irreversible changes, respectively, within 21 days of application of the substance (Section 3.3.1.1 of Annex I to the CLP Regulation). See also Sections 3.2 and 3.3 of the Guidance on the application of the CLP criteria.

3.3 Relevant test methods

The adopted test methods for skin corrosion/irritation are the following:

In vitro

Skin irritation:
- Skin Irritation Reconstructed Human Epidermis (RHE) Test Method (EU B.46/OECD TG 439).

Skin corrosion:
- Skin Corrosion Reconstructed Human Epidermis (RHE) Test Method (EU B.40 bis/OECD TG 431).
- Transcutaneous Electrical Resistance (TER) Test Method (EU B.40/OECD TG 430).
- *In Vitro* Membrane Barrier Test Method (OECD TG 435).

In vivo

- Acute Dermal Irritation/Corrosion (EU B.4/OECD TG 404).

The adopted test methods for serious eye damage/eye irritation are the following:

In vitro

- Bovine Corneal Opacity and Permeability (BCOP) Test Method (EU B.47/OECD TG 437).
- Isolated Chicken Eye (ICE) Test Method (B.48/OECD 438).
- Fluorescein Leakage (FL) Test Method (OECD TG 460).
- Short Time Exposure (STE) Test Method (OECD TG 491).
- Reconstructed human Cornea-like Epithelium (RhCE) Test Method (OECD TG 492).

In vivo

- Acute Eye Irritation/Corrosion (EU B.5/OECD TG 405).

The OECD also recently published a Guidance Document on serious eye damage and eye irritation IATA [36], which provides examples of how to combine multiple *in vitro* methods to improve the predictivity in identifying CLP Category 2 eye irritants.

For further details, see Section R.7.2 of the Guidance on IR&CSA – Chapter R.7a, and ECHA Advice on skin and eye irritation testing to help reduce animal tests.

3.4 Specific adaptation rules

Specific rules for adapting the information requirements for skin corrosion/irritation and serious eye damage/eye irritation are only mentioned in the REACH Regulation, not in the BPR (for further details see also Section R.7.2 of the Guidance on IR&CSA – Chapter R.7a).

In both REACH Annexes VII and VIII, column 2 adaptation rules for skin corrosion/irritation are based on physicochemical properties (i.e. pH, flammability) of the substance, its prior classification for acute toxicity by the dermal route (CLP Category 1), or the absence of skin irritation up to the limit dose level in an acute toxicity study by the dermal route.

In addition, the two *in vitro* studies required, i.e. one study for skin corrosion and one for skin irritation, don't need to be systematically performed if one is enough, i.e. results from one of them already allow a conclusive decision on the classification of a substance or on the absence of skin irritation potential.

Similarly, the *in vitro* and *in vivo* studies on serious eye damage/eye irritation do not need to

be conducted if the substance is a strong acid or base, it is spontaneously flammable in air or in contact with water or moisture at room temperature, or it is already classified as skin corrosive (CLP Category 1), leading to classification for serious eye damage (CLP Category 1), or skin irritant (CLP Category 2) and the available information indicate that the substance should be classified for eye irritation (CLP Category 2).

4. Skin and respiratory sensitisation

4.1 Information requirements

4.1.1 Under the REACH Regulation

REACH information requirements for skin sensitisation are specified in recently revised Section 8.3 of Annex VII (published in September 2016) and ask for information allowing the identification of skin sensitisers, including their potential classification in sub-category 1A to be concluded, and for performing a risk assessment, as required (see also Section R.7.3 of the Guidance on IR&CSA – Chapter R.7a).

More specifically, this information should come from:

- *in vitro/in chemico* data addressing each of the following three KEs of skin sensitisation adverse outcome pathway (AOP): molecular interaction with skin proteins, inflammatory response in keratinocytes, activation of dendritic cells (Point 8.3.1 of REACH Annex VII);
- an *in vivo* study for skin sensitisation (Point 8.3.2 of REACH Annex VII), normally a Local Lymph Node Assay (LLNA), if the above *in vitro/in chemico* studies are not applicable for the substance or are not adequate for classification and risk assessment.

Respiratory sensitisation is not a standard information requirement under the REACH Regulation. However, if data are available, they should be included in the technical dossier and used to support classification and labelling where relevant (see also Section R.7.3 of the Guidance on IR&CSA – Chapter R.7a).

4.1.2 For biocide active substances under the BPR

The information requirement for skin sensitisation is part of the core dataset and is divided in two steps as specified in Section 8.3 of BPR Annex II (see also Section 8.3 of the Guidance on BPR Volume III - Part A).

- Step 1: an assessment of the available human, animal and alternative data;
- Step 2: *in vivo* testing (LLNA is the first choice of assay when new information needs to be generated, justification for using another test needs to be provided).

The information on respiratory sensitisation is part of the additional dataset. The assessment of the potential of a substance to induce respiratory sensitisation should include assessment of the available existing information: non-human data (e.g. physicochemical properties, grouping, (Q)SARs and expert systems, *in vitro* data), human data and animal data.

4.2 CLP criteria

Hazard category and sub-categories for skin and respiratory sensitisers are based on the frequency of occurrence in humans and/or potency in animals (see Section 3.4 of the Guidance on the application of the CLP criteria).

4.3 Relevant test methods

The adopted test methods for skin sensitisation are as follows:

In vitro/in chemico

- Direct Peptide Reactivity Assay (DPRA) (EU B.59/OECD TG 442C) for KE1.
- ARE-Nrf2 Luciferase Test Method (EU B.60/OECD TG 442D) for KE2
- human Cell Line Activation Test (h-CLAT), U-SENS and IL-8 Luc Assay (OECD TG 442E) for KE3.

In vivo

- Local Lymph Node Assay (LLNA) (EU B.46/OECD 429).
- *In Vivo* Guinea Pig Test Method (EU B.6/OECD TG 406).
- Local Lymph Node Assay: DA (EU B.50/OECD TG 442A).
- Local Lymph Node Assay: BrdU-ELISA (EU B.51/OECD TG 442B).

The above adopted *in vitro/in chemico* test methods are based on the KEs as specified in the AOP for skin sensitisation initiated by covalent binding to proteins [119]: DPRA for KE 1, ARE-Nrf2 Luciferase Test Method for KE 2 and h-CLAT, U-SENS and IL-8 Luc Assay for KE 3.

Regarding the adopted *in vivo* test methods, it should be noted that the LLNA is the first choice *in vivo* assay and generation of new data with the *in vivo* Guinea Pig test method needs to be justified. In addition, the LLNA: DA and LLNA: BrdU-ELISA are not recommended for new testing as there are currently no CLP criteria available for predicting skin sensitisation potency with these methods, even if the dose-response relationship information obtained may provide some information that can be used within a WoE approach.

For further details on the recommended use of the above methods, please see Section R.7.3 of the Guidance on IR&CSA – Chapter R.7a and ECHA Advice on skin sensitisation testing to help reduce animal tests.

There are currently no standard tests and no adopted test methods available for respiratory sensitisation.

4.4 Specific adaptation rules

Under the REACH Regulation, the studies under Points 8.3.1 and 8.3.2 of REACH Annex VII do not need to be conducted if the substance is classified as skin corrosive (CLP Category 1), or it is a strong acid or base, or it is spontaneously flammable in air or in contact with water or moisture at room temperature (column 2 of REACH Annex VII, Point 8.3).

In addition, column 2 adaptation rules for Point 8.3.1 specifies that *in vitro/in chemico* tests do not need to be conducted if an *in vivo* study for skin sensitisation is available, or the available *in vitro/in chemico* test methods are not applicable to the substance or are not adequate for skin sensitisation classification and risk assessment.

Also, all three KEs indicated in Point 8.3.1, corresponding to the first three KEs of the OECD AOP, do not have to be addressed if classification and risk assessment for skin sensitisation is achieved with information on one or two of the KEs.

Regarding the need for *in vivo* skin sensitisation studies, those that have used EU or OECD-adopted and GLP-compliant test methods that were carried out or initiated before 10 May 2017, these are considered appropriate to address this standard information requirement.

Under the BPR, skin sensitisation testing is not needed if the available information indicates that the substance should be classified for skin sensitisation or corrosivity, or that it is a strong acid or base.

5. Repeated-dose and chronic toxicity

5.1 Information requirements

5.1.1 Under the REACH Regulation

The standard information requirements for repeated-dose toxicity are *in vivo* studies of increasing minimum duration as the tonnage band is higher. The oral route is the most common route of administration, but the choice of the route of administration depends on the substance properties and on the relevant exposure route for humans (see the REACH Regulation and Section R.7.5 of the Guidance on IR&CSA – Chapter R.7a).

Information on a sub-acute (28-day) study is needed at Annex VIII (10-100 tonnes per year) level. At the next tonnage band, a longer study, i.e. sub-chronic (90-day) study, is required. The rat is the standard species under the REACH Regulation.

In addition, further studies may be needed under Annexes IX and X to address concerns related to longer exposure duration, different routes of administration and/or specific toxicological investigations, such as immunotoxicity or neurotoxicity.

Long-term chronic toxicity studies may be needed based on human exposure considerations (see the REACH Regulation and Section R.7.5 of the Guidance on IR&CSA – Chapter R.7a).

5.1.2 For biocide active substances under the BPR

The core information requirements for repeated-dose toxicity and triggers for further studies under the BPR are basically the same as those under the REACH Regulation. The BPR also contains specific provisions for further testing in a second (non-rodent) species if justified.

5.2 CLP criteria

Categories for specific target organ-toxicity – repeated exposure are based on evidence from humans and/or from experimental animals and dependent on nature and severity of effects and effective dose levels (see Section 3.9 of the Guidance on the application of the CLP criteria). Classification based on gonad toxicity follows criteria for reproductive toxicity.

5.3 Relevant test methods

The relevant standard test methods for repeated-dose toxicity testing are listed below. The current standard test methods are all *in vivo* studies.

28-day repeated-dose toxicity studies

- Repeated-Dose 28-day Oral Toxicity Study in Rodents (EU B.7/OECD TG 407).
- Repeated-Dose Dermal Toxicity: 21/28-day Study (EU B.9/OECD TG 410).
- Sub-Acute Inhalation Toxicity: 28-Day Study (EU B.8/OECD TG 412).
- Combined Repeated-Dose Toxicity Study with the Reproduction/Developmental Toxicity Screening Test (OECD TG 422).

90-day repeated-dose toxicity studies

- Repeated-Dose 90-day Oral Toxicity Study in Rodents (EU B.26/OECD TG 408).
- Repeated-Dose 90-day Oral Toxicity Study in Non-Rodents (EU B.27/OECD TG 409).
- Sub-Chronic Dermal Toxicity: 90-day Study (EU B.28/OECD TG 411).
- Sub-Chronic Inhalation Toxicity: 90-day Study (EU B.29/OECD TG 413).

Chronic toxicity studies

- Chronic Toxicity Test (EU B.30/OECD TG 452).
- Combined Chronic Toxicity/Carcinogenicity Studies (EU B.33/OECD TG 453).

OECD TG 408 is currently being revised. The latest update of a test guideline (OECD TG and/or EU method) should always be used. Further details of the study protocols are described in the respective test guidelines and recommendations on their use can be found in Section R.7.5 of the Guidance on IR&CSA - Chapter R.7a and the OECD test guideline web page.

In addition, other types of studies with repeated administration and investigations on organ and tissue toxicity can provide information on repeated-dose toxicity, e.g. neurotoxicity studies, reproductive toxicity studies and carcinogenicity studies (See Section R.7.5 of the Guidance on IR&CSA - Chapter R.7a).

At OECD level, the following activities are ongoing:

- the Inhalation TGs and GD amended to Accommodate Nanomaterial Safety Testing have been approved in April 2017; and
- update of the repeated-dose oral toxicity 90-day study (OECD TG 408) with parameters for endocrine activity (see OECD Work plan for the Test Guidelines Programme 2017 [22]).

5.4 Specific adaptation rules

The repeated-dose toxicity study does not need to be conducted in certain cases specified in column 2 of Section 8.6 of REACH Annexes VIII to X and column 3 of Section 8.9 of BPR Annex II.

These waiving possibilities are based amongst others on the availability of existing data from reliable and appropriate *in vivo* studies, on substance classification for specific target organ toxicity after repeated exposure (STOT RE 1 or STOT RE 2) under the CLP Regulation, on substance transformation, reactivity and absorption data, and on human exposure data (for further details, see Section R.7.5 of the Guidance on IR&CSA - Chapter R.7a and Section 8.9 of the Guidance on BPR Volume III - Part A.

6. Mutagenicity

6.1 Information requirements

6.1.1 Under the REACH Regulation

The assessment of mutagenicity under the REACH Regulation follows a stepwise approach, starting with a battery of *in vitro* tests followed up by appropriate *in vivo* testing if one or more of the *in vitro* tests is positive.

The *in vitro* studies for mutagenicity include an *in vitro* gene mutation study in bacteria (Ames test) (Annex VII, Section 8.4.1), an *in vitro* cytogenicity study in mammalian cells (i.e. an *in vitro* chromosome aberration study or an *in vitro* micronucleus study) (Annex VIII, Section 8.4.2) and, if both first *in vitro* tests are negative, an *in vitro* gene mutation study in mammalian cells (Annex VIII, Section 8.4.3).

If there is a positive result in any of the above *in vitro* studies and there are no results available from an appropriate *in vivo* study already, an appropriate follow-up *in vivo* study in somatic cells must be proposed by the registrant (Annex IX, Section 8.4). Appropriate means here that the type of *in vivo* mutagenicity testing should reflect the type of *in vitro* mutagenicity test results observed (*i.e.* gene mutation, structural or numerical chromosome aberration). In some cases, a second *in vivo* somatic cell test (Annex X, Section 8.4) and/or investigation of the germ cell mutagenicity potential (Annexes IX-X, Section 8.4) may be needed.

As for any other endpoint under the REACH Regulation, the information required for a substance depends on its volume of production or importation (i.e. in general the higher the annual tonnage, the more data/studies are required).

For mutagenicity/genotoxicity, however, the rules set out in Annexes VII to X may in some cases require certain tests to be undertaken earlier than or in addition to the tonnage-triggered requirements.

For further details on the information requirements under the REACH Regulation, see Section R.7.7 of the Guidance on IR&CSA – Chapter R.7a.

6.1.2 For biocide active substances under the BPR

Under the BPR, *in vitro* mutagenicity assessment of active substances is part of the core dataset (Section 8.5 of BPR Annex II) whereas an *in vivo* genotoxicity study is part of the additional dataset (Section 8.6 of BPR Annex II). The information required for biocide active substances is basically the same as that required under the REACH Regulation and follows a similar tiered approach:

- collection and evaluation of existing genotoxicity data, including available *in vivo* data;
- generation of *in vitro* data using the standard *in vitro* test battery (i.e. an *in vitro* gene mutation test in bacteria, an *in vitro* cytogenicity test in mammalian cells and an *in vitro* gene mutation test in mammalian cells;
- an appropriate follow-up *in vivo* study in somatic cells, if there is a positive result in any of the above *in vitro* studies and there are no results available from an appropriate *in vivo* study (a second *in vivo* study in somatic cells may be necessary in certain cases);
- investigation of the germ cell mutagenicity potential if there is a positive result from an *in vivo* somatic cell study available, on the basis of all available data, including toxicokinetic evidence to demonstrate that the substance reached the tested organ.

For further details on the information requirements under the BPR, see Sections 8.5 and 8.6 of the Guidance on BPR Volume III – Part A.

6.2 CLP criteria

Hazard categories for germ cell mutagens are largely based on evidence on humans and/or *in vivo* mutagenicity tests in mammals (see Section 3.5 of the Guidance on the application of the CLP criteria).

6.3 Relevant test methods

The relevant standard test methods for mutagenicity/genotoxicity testing are listed below.

In vitro

- Bacterial reverse mutation test (EU B.13/14/OECD TG 471).
- *In vitro* mammalian cell gene mutation tests using the *Hprt* and *xprt* genes (EU B.17/OECD TG 476).
- *In vitro* mammalian cell gene mutation tests using the thymidine kinase gene (EU B.67/OECD TG 490).
- *In vitro* mammalian chromosome aberration test (EU B.10/OECD TG 473).
- *In vitro* micronucleus test (EU B.49/OECD TG 487).

In vivo

using somatic cells:

- *In vivo* mammalian bone marrow chromosome aberration test (EU B.11/OECD TG 475).
- *In vivo* mammalian erythrocyte micronucleus test (EU B.12/OECD TG 474).
- Unscheduled DNA synthesis (UDS) test with mammalian liver cells *in vivo* (EU B.39/OECD TG 486).
- Transgenic rodent (TGR) somatic and germ cell gene mutation assays (EU B.58/OECD TG 488).

- *In vivo* alkaline single-cell gel electrophoresis assay for DNA strand breaks (comet assay) (EU B.62/OECD TG 489).

using germ cells:
- Mammalian spermatogonial chromosome aberration test (EU B.23/OECD TG 483).
- Rodent dominant lethal test (EU B.22/OECD TG 478).
- Transgenic rodent (TGR) somatic and germ cell gene mutation assays (EU B.58/OECD TG 488).

Several of the above OECD test guidelines for mutagenicity/genotoxicity have recently been updated, but those changes have not yet been implemented in the EU Test Methods Regulation (TMR). As alignment of the test guidelines of the EU TMR with updated OECD test guidelines requires some time, the latest update of a test guideline (OECD TG and/or EU method) should be used for conducting new tests.

Further details of the study protocols are described in the respective test guidelines and recommendations on their use can be found in Section R.7.7 of the Guidance on IR&CSA - Chapter R.7a, ECHA's web page "Testing methods and alternatives" and the Overview of the set of OECD Genetic Toxicology Test Guidelines and updates performed in 2014-2015 [25].

6.4 Specific adaptation rules

These possibilities refer to situations where no test or information for a specific requirement on mutagenicity needs to be provided.

In vitro studies
Under the REACH Regulation, the first *in vitro* test in mammalian cells (*in vitro* cytogenicity study) and the second *in vitro* study in mammalian cells (*in vitro* mammalian cell gene mutation test) can normally be adapted if adequate information is available from *in vivo* studies addressing the same endpoints, i.e. cytogenicity [REACH Annex VIII, Column 2, Section 8.4.2] and gene mutation [REACH Annex VIII, Column 2, Section 8.4.3], respectively.

The first *in vitro* test in mammalian cells does not need to be conducted if the substance is classified in the most severe categories for carcinogenicity (i.e. Carc 1A or 1B), or germ cell mutagenicity (i.e. Muta 1A, 1B or 2) under the CLP Regulation [REACH Annex VIII, Column 2, Section 8.4.2]. There are no such adaptations under the BPR and all three *in vitro* tests are part of the core dataset.

In vivo studies
The standard information requirements of REACH Annex IX, Section 8.4 and BPR Annex II, Section 8.6 must be fulfilled if there is a positive result in any of the *in vitro* tests required under REACH Annexes VII and VIII and BPR Annex II, Section 8.5. This means that normally no *in vivo* tests will be required if all the required *in vitro* tests have given reliable and conclusive negative results.

In addition, *in vivo* genotoxicity testing is generally not needed under the BPR if the substance is known to be carcinogenic (i.e. classified as Carc. 1A or 1B under the CLP Regulation) or mutagenic (i.e. classified as Muta. 1A, 1B or 2 under the CLP Regulation) or if valid *in vivo* micronucleus data are generated within a repeated-dose toxicity study and the *in vivo* micronucleus test is the appropriate test to be conducted to address the information requirement for mutagenicity.

Consideration of the need for germ cell mutagenicity study
The basic principle is that a substance considered as a positive *in vivo* somatic cell mutagen

should also be considered as a possible germ cell mutagen unless data can be provided to the contrary. The considerations should be documented. Based on all available data, including toxicokinetic and toxicodynamic properties, WoE determination using expert judgement is needed to consider whether there is sufficient information to conclude that such a substance also poses a mutagenic hazard to germ cells or whether additional investigation is necessary (for detailed information on the criteria for classification of substances for germ cell mutagenicity under the CLP Regulation, see Section 3.5 of the Guidance on the Application of the CLP Criteria).

7. Carcinogenicity

7.1 Information requirements

Repeated-dose toxicity (see Section B.5 on Repeated-dose toxicity) studies are normally required (or triggered) at lower tonnage levels than that relevant for triggering a carcinogenicity study under the REACH Regulation or are requested as part of the BPR core dataset.

In the same way that positive *in vivo* genotoxicity tests may indicate a genotoxic carcinogenicity potential, repeated-dose toxicity may be informative about a possible carcinogenic potential by a non-genotoxic mechanism if hyperplasia or other pre-neoplastic effects are observed. These observations can be used to identify specific target tissues, inform about potential modes of action (MoAs) underlying the carcinogenic effect and assist in the development of dose-effect relationships [229].

7.1.1 Under the REACH Regulation

A carcinogenicity study is required for substances produced or imported at 1 000 tonnes per year or higher (Section 8.9.1 of REACH Annex X) if there is high human exposure to the substance (i.e. widespread dispersive use or frequent or long-term human exposure), and if the substance is classified for mutagenicity (germ cell mutagen category 2 under the CLP Regulation) or there is evidence from the repeated-dose studies that the substance is able to induce hyperplasia and/or pre-neoplastic lesions.

Moreover, the REACH Regulation also makes it possible for carcinogenic substances at all tonnage levels to be identified as substances of very high concern (SVHCs) to be included in the Candidate List for authorisation, taking into account information from all available relevant sources.

For further details, see Section R.7.7.9 of the Guidance on IR&CSA – Chapter R.7a.

7.1.2 For biocide active substances under the BPR

Under the BPR, a carcinogenicity study is part of the core dataset (Section 8.11 of BPR Annex II). If a new study is required, a combined chronic toxicity/carcinogenicity study in the rat and preferably through the oral route must be performed (Section 8.11.1 of BPR Annex II).

Additionally, a carcinogenicity study in the mouse is normally performed (Section 8.11.2 of BPR Annex II), although the need for this second study should be considered on a case-by-case basis.

For biocide active substances, a carcinogenicity study in a second species, normally the mouse, should be conducted. Case-by-case considerations on species-specificity (organ specificity of effects) and species differences and human relevance are needed when the need of information from the second species is considered.

Some publications suggest that a second study in another rodent species (than rats) is not likely to provide additional information [230]. The second test would, in any case, not be needed if the substance can be classified as germ cell mutagen category 1A or 1B under the CLP Regulation based on the first study.

For further details, see Section 8.11 of the Guidance on BPR Volume III - Part A.

7.2 CLP criteria

Hazard categories for carcinogens are largely based on human and/or animal evidence and strength of evidence (see Section 3.6 of the Guidance on the application of the CLP criteria).

7.3 Relevant test methods

- Carcinogenicity Study (EU B.32/OECD TG 451).
- Combined Chronic Toxicity/Carcinogenicity studies (EU B.33/OECD TG 453).

Although not aimed at investigating carcinogenicity *per se*, other available OECD/EU test guideline studies may provide useful information that can be indicative of a carcinogenic potential and determine the need for follow-up studies:

- Genotoxicity Studies: see Section B.6 on mutagenicity.
- Repeated-Dose Toxicity Studies: see Section B.5 on repeated-dose toxicity.

7.4 Specific adaptation rules

Under both the REACH (Annex X, Section 8.9.1, column 2) and BPR (Annex II, Section 8.11, column 3), a carcinogenicity study does not need to be conducted if the substance is classified under the CLP Regulation as germ cell mutagen category 1A or 1B.

In such a case, the default presumption would be that a genotoxic mechanism for carcinogenicity is likely and a carcinogenicity study will normally not be required.

In addition, for biocide active substances, the carcinogenicity study does not need to be conducted if:

- no genotoxic potential for humans is identified in genotoxicity tests; and
- possible mechanisms of toxicological effects observed in sub-chronic toxicity studies do not point at non-genotoxic carcinogenicity and there are no structural alerts for carcinogenicity; and
- the sub-chronic studies in rodents and/or non-rodents do not show any substance-related adverse effects at the limit dose level.

For further information for biocide active substances, see Section 8.11 of the Guidance on BPR Volume III – Part A.

8. Reproductive toxicity

8.1 Information requirements

8.1.1 Under the REACH Regulation

Under REACH, reproductive toxicity properties of a substance are characterised by investigations according to three different studies: a reproduction/developmental toxicity screening test, prenatal developmental toxicity studies in two species, and an extended one-

generation reproductive toxicity study.

In exceptional cases, information from an extended one-generation reproductive toxicity study in a second species or strain may be needed.

These cases may include situations where information on another species or strain seems to be more relevant than information from an existing study, or where the dose levels used in the first study are not adequate for classification and labelling and/or risk assessment. A two-generation reproductive toxicity study is sufficient to cover the standard information requirement (Column 1), instead of an extended one-generation reproductive toxicity study, if initiated before March 13, 2015. For further details, see Section R.7.6 of the Guidance on IR&CSA – Chapter R.7a.

The information required for a substance depends on the annual tonnage level of manufacturing/import or concern. If there is a specific concern, more information may be required on functional fertility of the offspring, developmental neurotoxicity and/or developmental immunotoxicity, which can be achieved by including the relevant expansion/ cohorts in an extended one-generation reproductive toxicity study. Further concerns, beyond those addressed using the (standard) information described above, can be addressed under substance evaluation using concern and risk-based argumentation, if it considered that more information is needed for adequate risk management actions.

The screening study for developmental and reproductive toxicity (OECD TGs 421/422) provides limited information on reproductive and developmental toxicity due to its more limited investigations and lower number of animals used per dose group than those of the extended one-generation reproductive toxicity study (EOGRTS, EU B.56/OECD TG 443) or the two-generation reproductive toxicity study (EU B.35/OECD TG 416).

The screening study therefore has a more limited detection power for developmental or reproductive toxic effects compared to these higher-tier studies. Reproduction of parental animals is investigated after exposing them shortly (two weeks) before mating, during pregnancy and lactation up to lactation day 13. The test method has recently been updated to include a few parameters sensitive to certain endocrine modes of action (i.e. antiandrogenicity and thyroidal effects). It is not an alternative or a replacement for the other test methods such as EU B.31/OECD TG 414 and EU B.56/OECD TG 443. It can be used as a range-finder for EU B.56/OECD TG 443.

The prenatal developmental toxicity study (EU B.31/OECD TG 414) provides information on developmental toxicity (lethality; growth; gross, visceral and skeletal malformations and variations) after exposure during *in utero* development (prenatal period). Usually, information from two species is considered sufficiently comprehensive for covering the uncertainties related to species-specific effects.

The extended one-generation reproductive toxicity study (EU B.56/OECD TG 344) is considered to provide more comprehensive information on reproductive toxicity than the screening for reproductive/developmental toxicity, in particular due to more life stages covered, leading to more in depth information. Furthermore, many more parameters are investigated and the detection power is much higher. The design of the extended one-generation reproductive toxicity study is specified under the REACH Regulation and further elaborated in Section R.7.6 of the Guidance on IR&CSA – Chapter R.7a. The basic study design is on effects on reproduction in parental animals and their offspring until adulthood. If there is a specified concern and triggers are met, more information may be required on sexual function and fertility of the offspring, developmental neurotoxicity and/or developmental immunotoxicity.

All the studies mentioned above – reproduction/developmental toxicity screening test, prenatal

developmental toxicity studies, extended one-generation reproductive toxicity study, two-generation reproductive toxicity study (initiated before 13 March 2015) – are, if conducted following EU and OECD test methods and endpoint specific REACH requirements, suitable for both risk assessment and classification and labelling purposes, including sub-categorisation.

8.1.2 For biocide active substances under the BPR

Regarding biocide active substances, information from a two-generation reproductive toxicity study and a prenatal developmental toxicity study in one species (the rabbit is preferred) are the main information sources.

If another reproductive toxicity test is used instead of a two-generation reproductive toxicity study, a justification must be provided why another study is more appropriate.

An extended one-generation reproductive toxicity study must be considered as an alternative approach to the multi-generation study. For reproductive toxicity, if there is a specific concern, the core dataset may be extended with an additional set of data, including a prenatal developmental toxicity study in a second species or a mechanistic study investigating mechanisms/modes of action, information on developmental neurotoxicity, developmental immunotoxicity and studies evaluating endocrine disruption. For further details, see Section 8.10 of the Guidance on BPR Volume III – Part A.

8.2 CLP criteria

CLP criteria for hazard categories for reproductive toxicants are largely based on strength of evidence from humans and/or data from animal studies (see Section 3.7 of the Guidance on the application of the CLP criteria).

8.3 Relevant test methods

- Reproduction/developmental toxicity screening test (OECD TG 421).
- Combined repeated-dose toxicity study with the reproduction/developmental toxicity screening test (OECD TG 422).
- Prenatal developmental toxicity study (EU B.31/OECD TG 414).
- Extended one-generation reproductive toxicity study (EU B.56/OECD TG 443)(REACH and BPR).
- Two-generation reproductive toxicity study (EU B.35/OECD TG 416)(BPR).

Repeated-dose toxicity studies may provide information on effects on gonads and accessory sex organs (organ weight changes and histopathology).

Test methods specifically investigating potential endocrine modes of action (MoAs) are presented in a separate section. However, the extended one-generation reproductive toxicity study also includes parameters measuring potential endocrine MoAs and sexual maturation (e.g. oestrous cycle, vaginal opening, time from vaginal opening to first oestrous cycle, preputial separation, thyroid hormone levels, nipple areola retention, anogenital distance).

Also the two-generation reproductive toxicity study includes investigations for endocrine MoAs and sexual maturation, but a few less or less frequent measurements than those of the extended one-generation reproductive toxicity study (oestrous cycle, vaginal opening, preputial separation, anogenital distance if triggered). OECD TGs 421 and 422 have been recently updated with investigations reflecting certain endocrine MoAs (thyroid hormone measurements, nipple/areola retention, anogenital distance).

Current OECD activities for this endpoint include:

- update of Guidance Document 150 (Guidance document on standardised test guidelines for evaluating chemicals for endocrine disruption);
- minor enhancements of OECD TG 414 (Prenatal Developmental Toxicity Study) with ED-relevant endpoints (thyroid hormone measurements, anogenital distance), based on feasibility study.

8.4 Specific adaptation rules

The possibilities to apply adaption rules refer to situations where no new test or information on reproductive toxicity needs to be provided based on applicability of relevant column 2 adaptations (specific adaptation rules) in Section 8.7 of REACH Annexes VIII-X, due to technical reasons according to REACH Annex XI, Section 2 or when human exposure can be excluded following rules of REACH Annex XI, Section 3 (see also Section R.7.6 of the Guidance on IR&CSA - Chapter R.7a). These adaptations are basically similar under the BPR (specific adaptation rules of column 3 in Section 8.10 of Annex II and general adaptation rules of BPR Annex IV) and described in Section 8.10 of the Guidance on BPR Volume III – Part A.

The idea behind these adaptations is that if the safe use of a substance is already managed by restricting its use and exposure due to other hazardous properties, such as germ cell mutagenicity or genotoxic carcinogenicity, then information on reproductive toxicity will not change the risk management measures necessary and thus, further information on reproductive toxicity is not needed. Similarly, if there is no exposure to the substance, no information on reproductive toxicity is needed.

The following adaption rules are applicable when the information requirement is normally a reproduction/developmental toxicity screening test (OECD TG 421/422), prenatal developmental toxicity studies (EU B.31/OECD TG 414) and extended one-generation reproductive toxicity study (EU B.56/OECD TG 443)/two-generation reproductive toxicity study (EU B.35/OECD TG 416):

- Information may not be needed if the substance is classified in one of the most severe categories for carcinogenicity or mutagenicity and appropriate risk management measures are in place.
    - In practice, this means classification as Carc. 1A or 1B, Muta. 1A or 1B.
- Similarly, if the substance meets the criteria for classification in the most severe categories for reproductive toxicity and the data are adequate to support a robust risk assessment, information on reproductive toxicity may not be needed.
    - In practice, this means classification as Repr. 1A or 1B for both sexual function and fertility ("May damage fertility"; H360F) and developmental toxicity ("May damage the unborn child"; H360D) [REACH Annexes VIII, IX and X, Column 2, Section 8.7; BPR Annex II, Column 3, Section 8.10].
    - If a substance is classified as Repr. 1A or 1B for sexual function and fertility, information is still needed on developmental toxicity.
    - The same applies if a substance is classified as Repr. 1A or 1B for developmental toxicity. In those cases, information on sexual function and fertility is still needed, because this information could result in different derived no-effect levels (DNELs) and more importantly this could result in complete different outcomes in the health impact assessment.

In addition to the above, the following is applicable when the information requirement is normally a reproduction/developmental toxicity screening test (OECD TG 421/422):

- The information requirement may be adapted in REACH Annex VIII if a prenatal developmental toxicity study, or either an extended one-generation reproductive toxicity study or a two-generation reproductive toxicity study is available [REACH Annex VIII, Column 2, Section 8.7].

- The information requirement may also be adapted if relevant human exposure can be excluded. There are three alternative sets of conditions that may apply for a reproduction/developmental toxicity screening test (OECD TG 421/422) based on exposure scenarios [REACH Annex XI, Section 3.2 (a), (b) and (c)].

The following adaption rules are applicable when the information requirements are normally a prenatal developmental toxicity study and an extended one-generation toxicity study/two-generation reproduction toxicity study:

- There is a possibility to adapt the study if the substance shows no toxicity, there is no absorption and no (or no significant) human exposure. This is a very strict rule because all three requirements need to be met. No toxicity is defined by "low toxicological activity" meaning no evidence of toxicity is seen in any of the tests available provided that the dataset is sufficiently comprehensive and informative. No absorption means that there are toxicokinetic data showing that the substance does not enter the body and into the bloodstream through relevant routes of exposure, that it cannot be detected by methods with an appropriate sensitivity and that the substance and its metabolites are absent in urine, bile and exhaled air. No significant human exposure is difficult to define, but it means a very low exposure, if any [REACH Annexes IX and X, Column 2, Section 8.7; BPR Annex II, Column 3, Section 8.10].

Information on prenatal developmental toxicity and/or extended one-generation reproductive toxicity or two-generation reproductive toxicity (BPR) may be adapted if relevant human exposure can be excluded by two alternative sets of conditions [REACH Annex XI, Sections 3.2 (b) and (c); for BPR see Annex IV, Section 3].

9. Neurotoxicity/developmental neurotoxicity

9.1 Information requirements

Information on neurotoxicity or developmental neurotoxicity, in addition to parameters included in repeated-dose toxicity studies and reproductive toxicity studies, is normally required only based on concern.

Standard toxicity studies may indicate neurotoxicity and then further investigations may be needed. Further investigation is possible using standard repeated-dose toxicity tests (such as 28- and 90-day repeated-dose toxicity studies or the extended one-generation reproductive toxicity study) with incorporation of specific neurotoxicity measurements.

Neurotoxicity studies in rodents, if available, normally provide sufficient data to evaluate the potential neurotoxicity of the substance (neurobehavioural and neuropathological effects) after single and repeated exposure.

9.1.1 Under the REACH Regulation

Specific information beyond the information provided by the *in vivo* standard test methods is not normally required for dossier evaluation under the REACH Regulation. However, based on a particular concern for neurotoxicity, specific studies in young adults may be conducted or, if the concern is on developmental neurotoxicity, a developmental neurotoxicity cohort can be included in the extended one-generation reproductive toxicity study, or a separate developmental neurotoxicity study may be appropriate (See Sections R.7.5 and R.7.6 of the Guidance on IR&CSA – Chapter R.7a).

The developmental neurotoxicity test method (OECD TG 426) is designed to be performed as an independent study. Specific studies, including *in vitro* studies or testing strategies, may be

requested under the substance evaluation process.

Developmental neurotoxicity investigations are triggered if the substance or a structurally similar substance has been shown to cause neurotoxicity in adults, structural abnormalities of the central nervous system in adults or in developing organism, or have a mode of action that has been closely linked to neurotoxic or developmental neurotoxicity effects e.g. cholinesterase inhibition or thyroid effects. Both results from animal studies and non-animal approaches are relevant. More details are provided in Section R.7.6 of the Guidance on IR&CSA – Chapter R.7a.

9.1.2 For biocide active substances under the BPR

The core dataset does not include specific information on neurotoxicity or developmental neurotoxicity beyond the parameters included in repeated-dose toxicity studies. However, if the mechanisms of action for an active substance is known (e.g. organophosphorous substances, carbamates, pyrethroids, etc.) or if there is any evidence from repeated-dose toxicity studies showing that the active substance may have neurotoxicity and developmental neurotoxic properties, then additional information or specific studies are needed.

The ECHA guidance specifies certain requirements for neurotoxicity investigations, e.g. for the route and species to be used (Section 8.13.2 of the Guidance on BPR Volume III – Part A).

In exceptional cases, investigations on developmental neurotoxicity are relevant. The substance may have been shown to:

- cause structural abnormalities of the central nervous system;
- cause clear signs of behavioural or functional adverse effects of nervous system involvement in adult animals; or
- have a mode of action that has been closely linked to neurotoxic or developmental neurotoxicity effects e.g. cholinesterase inhibition or thyroid effects.

However, in the case of (3), targeted testing on the specific mode of action in developing animals may provide sufficient information for regulatory purposes.

The developmental neurotoxicity study (EU B.53/OECD TG 426) is designed to be performed as an independent study and investigations can also be added to a two-generation reproduction study.

Developmental neurotoxicity investigations (Cohorts 2A and 2B) may also be included in the extended one-generation reproductive toxicity study (EU B.56/OECD TG 443). More details are provided in ECHA guidance (Section 8.13.2 of the Guidance on BPR Volume III – Part A).

9.2 CLP criteria

There are no specific CLP criteria for developmental neurotoxicity. Developmental neurotoxicity can be classified under reproductive toxicity (developmental toxicity). Adult neurotoxicity is classified under specific organ toxicity.

9.3 Relevant test methods

- Neurotoxicity study in rodents (EU B.43/OECD TG 424).
- Delayed neurotoxicity of organophosphorus substances after acute exposure (EU B.37/OECD TG 418).
- Delayed neurotoxicity of organophosphorus substances 28-day repeated-dose study
- Developmental neurotoxicity study (EU TM B.53/OECD TG 426).
- Extended one-generation reproductive toxicity study (EU B.56/OECD TG 443).

(EU TM B.38/OECD TG 419).

Neurotoxicity parameters can also be included in other repeated-dose toxicity studies. Investigations may include e.g. neuropathology, immunocytochemistry, use of special strains, electrophysiology, functional observations, sensory function tests, motor function tests, cognitive function tests, neurotransmitter analyses, enzyme/protein activity and measures of cell integrity.

When designing a (developmental) neurotoxicity study, it should be kept in mind that the areas of neurodevelopment and neurotoxicity are inherently very complex, and, in particular, there are massive gaps in knowledge about normal brain development at the functional, structural and molecular levels. This complicates both developmental neurotoxicity testing as well as substance safety and health risk assessment. Also, it makes it difficult to define strict criteria for testing, data interpretation, and risk assessment.

Therefore, investigation of (developmental) neurotoxicity should remain flexible to enable the design of the most sensitive and appropriate study relevant for the exposure and toxicity of the tested substance [204].

10. Immunotoxicity/developmental immunotoxicity

10.1 Information requirements

The need for further testing to characterise effects of concern for immunotoxicity has to be considered on a case-by-case basis. The conduct of the repeated-dose toxicity tests and the reproductive toxicity tests should be performed in a way that allows immunotoxicity potential to be evaluated to the extent possible. For example, an OECD TG 443 - extended one-generation reproductive toxicity study - may be conducted with the immunotoxicity cohort.

10.1.1 Under the REACH Regulation

Specific information outside the information provided by the *in vivo* test methods is not normally required. However, based on concern, immunotoxicity can be investigated in adults, e.g. by adding additional investigations into the repeated-dose toxicity study design (OECD TG 408 allows functional tests such as TDAR to be included into the TG or a developmental immunotoxicity cohort can be included in an extended one-generation reproductive toxicity study (Section R.7.6 of the Guidance on IR&CSA – Chapter R.7a).

Specific studies can also be required under the substance evaluation process.

10.1.2 For biocide active substances under the BPR

Core information requirements do not include specific information on immunotoxicity or developmental immunotoxicity in addition to parameters included in animal studies according to information requirements.

Additional information or specific studies on immunotoxicity, including developmental immunotoxicity, must however be required if there is any evidence, from skin sensitisation, repeated-dose or reproduction toxicity studies, that the active substance may have immunotoxic properties (Section 8.13.4 of the Guidance on BPR Volume III – Part A).

The information should elucidate the mechanism/mode of action and provide sufficient evidence for relevant adverse effects in humans.

The objectives of investigating immunotoxicity are to evaluate:

- whether the substance of interest has the potential to induce adverse effects involving the immune system; special attention should be paid to the adverse immunotoxic outcome among susceptible and vulnerable groups such as early and late life stages;
- the adverse outcomes caused by exposure to the substance (inflammation, immunosuppression; increased propensity for allergic disease; hypersensitivity reactions directed to the substance itself; increased risk of autoimmune disease; dysfunctional responses resulting in tissue or organ damage or dysfunction; impact on the developing immune system).

Guidance for the evaluation of all available information before conducting new tests is available in Section 1.7.3.4 of the Guidance on BPR Volume III – Parts B+C and is largely based on the WHO/IPCS Guidance on Immunotoxicity for Risk Assessment [207].

10.2 CLP criteria

There are no specific CLP criteria for developmental immunotoxicity. Developmental immunotoxicity can be classified under reproductive toxicity (developmental toxicity). Adult immunotoxicity is classified under specific organ toxicity.

10.3 Relevant test methods

Currently, there are not many internationally accepted standard test methods and therefore also only little test data with such methods. Tests for regulatory purpose include only that for skin sensitisation and immunosuppression (Cohort 3 in OECD TG 443). All the other methods have been used for research purposes and no internationally accepted standard test methods are yet available. However, information from these can be used especially under substance evaluation for other aspects such as for autoimmunity.

Different test methods can be employed for assessing immune suppression, immune stimulation and autoimmunity as well as developmental immunotoxicity.

It should also be noted that current animal studies provide information from an unchallenged immune system that has potential pitfalls in the assessment of immunotoxic potential (WHO/IPCS guidance for immunotoxicity risk assessment for substances [207]).

Reviews of principles and methods for immunotoxicity are available from WHO/IPCS:

- WHO/IPCS Environmental Health Criteria (EHC) 180, Principles and Methods for Assessing Direct Immunotoxicity Associated with Exposure to Chemicals (WHO, 1996 [231]).
- WHO/IPCS Environmental Health Criteria (EHC) 212, Principles and Methods for Assessing Allergic Hypersensitization Associated with Exposure to Chemicals (WHO, 1999 [232]).
- WHO/IPCS Environmental Health Criteria (EHC) 236, Principles and Methods for Assessing Autoimmunity Associated with Exposure to Chemicals (WHO, 2007 [233]).
- WHO/IPCS Guidance for immunotoxicity risk assessment for chemicals, Harmonisation project document No. 10 (WHO, 2012 [207]).

A list of methods that can be considered for immunotoxicity testing is provided below. This list is not exhaustive but provides the methodological aspects to consider on a case-by-case basis.

Immune suppression

- ICH Topic S 8 Immunotoxicity Studies for Human Pharmaceuticals (CHMP/167235/2004).
- US EPA OPPTS 870.7800 Health Effects Test Guidelines Immunotoxicity.
- Functional studies as described under Additional Immunotoxicity Studies below.

Immune stimulation
- Skin sensitisation (LLNA assay, see sensitisation section).
- Respiratory sensitisation (no adopted OECD test guidelines available).
- Autoimmunity (no adopted OECD test guidelines available).

Developmental immunotoxicity
- OECD TG 443: Extended One-Generation Reproductive Toxicity Study.

Additional Immunotoxicity Studies
- T-cell Dependent Antibody Response (TDAR).
- Immunophenotyping.
- Natural Killer Cell Activity Assays.
- Host Resistance Studies.
- Macrophage/Neutrophil Function.
- Assays to Measure Cell-Mediated Immunity.

11. Endocrine-disrupting properties

11.1 Information requirements

The REACH Regulation and the BPR do not require specifically information on endocrine-disrupting properties by default. Data on reproductive toxicology and organ toxicity might contain relevant information on these properties. Furthermore, available information that is not part of the dossier (both from standard and non-standard assays) can and should be included in the assessment.

Information on endocrine-disrupting properties should be evaluated according to the definition of the WHO/IPCS: "An endocrine disruptor is an exogenous substance or mixture that alters functions of the endocrine system and consequently causes adverse health effects in an intact organism, or its progeny, or (sub-)populations." [214], which means that the identification of an endocrine-disrupting substance should fulfil the following conditions: (i) has an endocrine MoA, (ii) provokes adverse effects and (iii) there is a plausible link between adverse effects and endocrine MoA. WHO/IPCS further makes the distinction between an endocrine disruptor and a potential endocrine disruptor, which "[…] is an exogenous substance or mixture that possesses properties that might be expected to lead to endocrine disruption in an intact organism, or its progeny, or (sub-)populations."

The available endpoints can be put in the context of the OECD Conceptual Framework (CF) for the Screening and Testing of Endocrine-Disrupting Chemicals which defines five levels of an assessment of endocrine-disrupting properties. Information on endocrine MoA can be obtained from tests giving information equivalent to OECD CF 1-5 (such as read-across to other substances, *in vitro* tests and *in vivo* screening tests), while information on endocrine mediated adverse effects can be obtained from tests giving information equivalent to OECD CF 4-5.

OECD TGs and Guidance Documents for the examination of endocrine-disrupting properties as well as the Guidance on this topic by the European Commission, ECHA and EFSA should be considered. Relevant test methods and screening approaches are listed in Section 11.3 below.

11.1.1 Under the REACH Regulation

There is currently no specific standard information requirements for endocrine-disrupting properties. However, some of the standard information requirements describe studies which investigate some endocrine-disrupting modes of action (MoAs) and may indicate adverse effects, e.g. extended one-generation reproductive toxicity study (Section R.7.6 of the Guidance on IR&CSA - Chapter R.7a). Furthermore, specific studies can be required under the

substance evaluation process if concerns on endocrine disruption have been substantiated. For ecotoxicology see also Section R.7.8-4 of the Guidance on IR&CSA - Chapter R.7b.

Such information may come from non-animal approaches indicating concern based on structural resemblance with substances with known endocrine-disrupting properties or due to predicted or measured activity in MIEs or KEs of various endocrine-disrupting-related adverse outcome patheays (AOPs) or modes of action (MoAs). A comprehensive animal study may provide information on both likely MoAs but also adverse effects may be needed to address concerns for endocrine-disrupting MoAs stemming from non-animal approaches e.g. under substance evaluation. Information on endocrine-disrupting activity and related adverse effects are needed when considering identification of a substance as an SVHC under REACH Art. 57(f) (see also SVHC Roadmap to 2020 [234]).

11.1.2 For biocide active substances under the BPR

Information on endocrine disruption is required as part of the additional dataset if there is any evidence from *in vitro*, repeated-dose or reproduction toxicity studies, that the active substance may have endocrine-disrupting properties. Specific studies are required to elucidate the mechanism/mode of action (MoA) and provide sufficient evidence for relevant adverse effects. The assessment of endocrine-disrupting properties is an important aspect within the active substance assessment according to the BPR as substances which are EDs fulfil the exclusion criteria according to Article 57 (1)(d).

Information derived from the use of expert systems that indicate structural similarities to known endocrine disruptors should be taken into account in deciding the need for additional testing. Expert judgement is needed to decide whether there is a need to perform additional tests or whether the existing information can be used to conclude that the substance is an endocrine disruptor. Currently, a guidance on identifying substances as endocrine disruptors is under development in the EU based on the criteria developed by the European Commission (see the European Commission Policy).

11.2 CLP criteria

There are no hazard classes for endocrine disruption or endocrine modes of action under the CLP Regulation. CLP classification is based on the hazard outcome (e.g. for reproductive toxicity or carcinogenicity) and not on the endocrine modes of action.

11.3 Relevant test methods

The most recent information on OECD TGs can be found on the OECD page on test methods: http://www.oecd.org/env/ehs/testing/oecdguidelinesforthetestingofchemicals.htm

Screening methods (not standard information requirements under the REACH Regulation or the BPR):

In vitro (anti)oestrogenicity
- Draft updated OECD TG 455 (Estrogen receptor transactivation assays; Performance-Based Test Guideline for Stably Transfected Transactivation In Vitro Assays to Detect Estrogen Receptor Agonists and Antagonists).
- OECD TG 493 (Performance-Based Test Guideline for Human Recombinant Estrogen Receptor (hrER) In Vitro Assays to Detect Chemicals with ER Binding Affinity).

In vitro (anti)androgenicity
- OECD TG 458 (Stably Transfected Transcriptional Human Androgen Receptor Transcriptional Activation Assay for Detection of Androgenic Agonist and Antagonist Activity of Chemicals).

In vitro steroidogenesis
- OECD TG 456 (H295R Steroidogenesis Assay).

In vitro thyroid toxicity
No specific OECD TG yet.

In vivo screening studies with intact or non-intact animals
- OECD TG 440 (Uterotrophic Bioassay in Rodents).
- OECD TG 441 (Hershberger Bioassay in Rats).

In vivo standard information studies with parameters sensitive to endocrine MoA
- OECD TG 407 (Repeated-Dose 28-day Oral Toxicity Study in Rodents).
- OECD TG 421 (Reproduction/Developmental Toxicity Screening Test).
- OECD TG 422 (Combined Repeated-Dose Toxicity Study with the Reproduction/Developmental Toxicity Screening Test).
- OECD TG 443 (Extended One-Generation Reproductive Toxicity Study).

Furthermore, all repeated-dose toxicity studies may provide some but normally rather limited information on endocrine-disrupting properties.

Ecotoxicological studies with parameters for endocrine MoAs
Various test may provide information on endocrine MoAs, especially reproduction and development tests and tests focusing on endocrine-disrupting properties, such as thyroid toxicity. Examples of ecotoxicology tests are:
- OECD TG 206 (Avian reproduction test).
- OECD TG 231 (Amphibian metamorphosis assay).
- OECD TG 229 (Fish short-term Reproduction assay (FSTRA)).
- OECD TG 230 (21-Day fish assay).
- Variant of OECD TG 230 (OECD GD 148 Androgenised female stickleback screen (AFSS)).
- OECD TG 234 (Fish sexual development test (FSDT)).
- OECD TG 240 (MEOGRT) Medaka Extended One Generation Reproduction Test.
- OECD TG 241 (LAGDA) Larval Amphibian Growth and Development Assay.

Other studies
Assays validated by the US EPA which are mentioned in OECD GD 150:
- US EPA OPPTS 890.1250 (ER binding assay).
- US EPA OPPTS 890.1150 (AR binding assay).
- US EPA OPPTS 890.1200 (Aromatase assay).
- US EPA OPPTS 890.1500 (Pubertal development and thyroid function assay in peripubertal male rats (Male PP assay).
- US EPA OPPTS 890.1450 (Pubertal development and thyroid function assay in peripubertal female rats (Female PP assay).
- US EPA OPPTS 850.1500 (Fish lifecycle toxicity test (FLCTT).

OECD GD 150
The OECD has published an OECD Conceptual Framework for Testing and Assessment of Endocrine Disruptors (as revised in 2012) which lists the OECD Test Guidelines and standardised test methods available, under development or proposed that can be used to evaluate substances for endocrine disruption.

The framework is intended to provide a guide to the tests available which can provide information for assessing endocrine disruptors but is not intended to be a testing strategy. Further information regarding the use and interpretation of these tests is available in OECD Guidance Document No. 150 (Guidance document on standardised test guidelines for evaluating chemicals for endocrine disruption) [215].

http://www.oecd.org/env/ehs/testing/OECD%20Conceptual%20Framework%20for%20Testing%20and%20Assessment%20of%20Endocrine%20Disrupters%20for%20the%20public%20website.pdf

Screening approaches
ECHA has developed a **screening approach** to (de)prioritise substances for regulatory processes, this approach includes also screening for potential endocrine-disrupting properties [235].

JRC has also developed a screening methodology to identify potential endocrine disruptors according to different options [236]. This was developed in the context of an impact assessment of proposed criteria for identifying endocrine disruptors and their implementation in EU legislation.

In many cases combination of various assays in a prediction models or tiered screening may be useful, when available, instead of relying on individual separate assays. One example of tiered approach is a published tiered high-throughput screening approach to identify thyroperoxidase inhibitors within the ToxCast Phase I and II Chemical Libraries [237]. However, assessing potential endocrine-disrupting properties/MoA is based on WoE evaluation of available information. Furthermore, negative result in a screening for instance for oestrogenicity does not mean that the substance does not have oestrogenic activity.

The current activities on test methods and guidance documents at OECD level include:

- New TG: performance-based test guideline on Androgen Receptor transactivation assays.
- Update of TG 455 - TG 457 to include the Transcriptional Estrogen Receptor alpha CALUX Assay for detecting (anti)oestrogenic chemicals.
- Development of a set of reference chemicals for testing *in vitro* metabolism systems in oestrogen-androgen-steroidogenesis assays.
- Androgen-Receptor Transactivation Assay for detecting substances with (anti)androgenic potential using 22Rv1/MMTV cells.
- Exploring the concept of developing pathway-based test method performance metrics: a case study using Estrogen Receptor signalling.
- Elaborating the conceptual framework for testing and assessment of endocrine disrupting chemicals for cross linkage between human and ecotoxicology components: three case studies to supplement GD 181.
- Validation of the Xenopus Embryonic Thyroid Signalling Assay (XETA) (relevant for 3R methods and screening approaches).
- Detailed review paper on retinoid system.

12. Bioaccumulation in fish

12.1 Information requirements

12.1.1 Under the REACH Regulation

REACH Annex IX indicates that information on bioaccumulation in aquatic species, preferably fish, is required for substances manufactured or imported in quantities of 100 tonnes per year or more.

As well as being used for risk assessment, evaluation of the bioaccumulation endpoint is related to the persistent, bioaccumulative and toxic (PBT)/very persistent, very bioaccumulative (vPvB) screening and assessment under REACH Annex XIII.

The PBT/vPvB assessment is required for all organic substances (including organo-metals) for which a chemical safety assessment (CSA) must be conducted and reported in the chemical safety report (CSR) (i.e. tonnage band >10 tonnes per year, according to Article 14(1) of the REACH Regulation). The PBT/vPvB assessment should also address the relevant constituents, impurities, additives and transformation/degradation products.

According to REACH Annex XIII, a substance fulfils the bioaccumulative criterion (B) when the bioconcentration factor (BCF) in aquatic species is greater than 2 000 and the very bioaccumulative criterion (vB) when the BCF is greater than 5 000.

12.1.2 For biocide active substances under the BPR

Regarding biocide active substances, information on bioconcentration (Annex II 9.1.4) can be provided based on estimation methods or on experimental determination and must be provided as part of the core dataset. The data requirement on bioconcentration is closely related to endpoint 9.1.7 on bioaccumulation which is considered an additional dataset and therefore not mandatory for all substances and only required under certain circumstances (related to the intrinsic properties of the substance to partition to lipids). An estimate of the bioconcentration factor related to absorption of the substance through the food chain should be provided for the evaluation of aquatic bioconcentration.

For biocides, an estimation of the intrinsic potential for bioconcentration in aquatic organisms should be submitted on the basis of physical and chemical properties (e.g. partition coefficient n-octanol/water). For surface active substances (surface tension lower than 60 mN/m) and dissociating or inorganic substances such as metals, toxicokinetic studies (including metabolism), residue studies or monitoring data on aquatic organisms (e.g. residue data in aquatic organisms and environmental concentrations) should be submitted.

It is mentioned in the Guidance on BPR Volume IV – Part A, Chapter II *Requirements for Active Substances* that *"[t]he experimental determination may not need to be carried out if it can be demonstrated on the basis of physico-chemical properties (e.g. log K_{ow} <3) or other evidence that the substance has a low potential for bioconcentration. All critical aspects of bioaccumulation such as ionic speciation, surface activity and metabolic transformation rates must be considered before experimental determination is considered unnecessary."*

Regarding experimental determination, the recommended test methods are OECD TG 305 or the EU method C.13 (Bioconcentration: Flow-through Fish Test) and specifically for marine environments, *Cyprinodon variegatus* should be tested.

Experimental testing is required when potential for bioaccumulation is indicated (e.g. by log K_{ow} ≥3, surface activity (i.e. surface tension <60 mN/m at a concentration <1 g/l), structural features indicating bioaccumulation (as in the case of e.g. pyridinium compounds), or risk for secondary poisoning).

12.2 CLP criteria

Under the CLP Regulation, a fish bioconcentration factor (BCF) is relevant for aquatic chronic classification in cases where a full set of long-term aquatic toxicity data is not available.

Bioaccumulation of a substance into an organism is not a hazard in itself, but should be considered as such, in relation to potential long-term effects. While, for organic substances, the potential for bioaccumulation may be determined by using the octanol/water partition coefficient, an experimentally determined BCF provides a better measure for the potential to bioaccumulate and is generally preferred, if available. For classification purposes, a BCF fish of ≥500 is indicative of the potential to bioconcentrate. All other provisions on bioaccumulation of

the REACH Regulation apply (see Section 4.1 of the Guidance on the application of the CLP criteria).

More details on the use of BCFs and potential non-animal approaches for CLP purposes can be found in Annex III to the Guidance on the Application of the CLP Criteria. In general, most principles from the REACH Guidance on IR&CSA – Chapter R.7c also apply for CLP with respect to this endpoint.

12.3 Relevant test methods

Fish tests

Risk assessments are preferably based on aquatic bioconcentration data (BCF). Similarly, the biaccumulative and very bioaccumulative criteria for the PBT/vPvB assessment are defined according to BCF values. However, other data can be used in a WoE approach such as the biomagnification factor (BMF), trophic magnification factor (TMF) and bioaccumulation factor (BAF) or biota-sediment/soil accumulation factor (BSAF) values. This is explained further in the Guidance on IR&CSA - Chapter R.7c.

In 2012, OECD TG 305 on bioaccumulation in fish was updated. Only the latest version of OECD TG 305 should now be used for generating new data with fish under the REACH Regulation. OECD TG 305 is used to assess the bioaccumulation potential of substances in fish.

In most cases, aqueous exposure is used, but dietary exposure is recommended for substances where the aqueous exposure methodology is not technically possible or feasible. The option of exposure solely through the dietary route generates a dietary BMF. The dietary BMF from the OECD TG 305 test differs from a field study's BMF value in which both water and dietary exposure may be combined. The dietary exposure route should be considered when it is impossible to maintain and measure aqueous concentrations of the test substance and/or when it is expected that exposure through the dietary route will be dominant, e.g. for substances with a high adsorption potential or low water solubility.

Approaches are available to estimate a kinetic bioconcentration factor (i.e. a BCF value) from data generated in the dietary study and these are discussed further in the test guideline and in OECD GD 264 [238].

OECD TG 305 also includes the option to perform a minimised aqueous exposure fish test. This method is described in paragraphs 83-96 of OECD TG 305. It uses a reduced number of sampling points and thus uses fewer fish. The conditions for selecting this option are described in the test guideline and further guidance is available in OECD GD 264 [238].

Another option in OECD TG 305 is to use one rather than two exposure concentrations. It therefore uses fewer fish than the two concentration test. Paragraphs 49-51 of the OECD TG 305 explain the conditions under which use of a single exposure concentration is possible and further guidance is available in OECD GD 264 [238].

More information on the conduct and interpretation of OECD TG 305 can be found in Section R.7.10 of the Guidance on IR&CSA - Chapter R.7c and in an OECD GD 264 [238].

12.4 Specific adaptation rules

The REACH standard information requirement on bioaccumulation in aquatic species can be adapted if it can be shown that the substance has a low potential for bioaccumulation (for

instance, a log K_{ow} ≤ 3) and/or a low potential to cross biological membranes, or direct and indirect exposure of the aquatic compartment is unlikely.

In addition, indicators of limited bioaccumulation potential include large molecular size, high log K_{ow} or low octanol solubility in combination with a lack of chronic toxicity in mammals and birds and no uptake in mammalian toxicokinetics studies. Further guidance is available in Section R.11.4.1.2 of the Guidance on IR&CSA - Chapter R.11.

When adapting the information requirement on bioaccumulation it should be carefully explained why the specific adaptation rules apply to the substance. For example, it should be considered if log K_{ow} is the appropriate measure to indicate low potential for bioaccumulation (see Section B.12.3 and Appendix R.7.10—3 Considerations for difficult substances in the Guidance on IR&CSA - Chapter R.7c).

Similarly for biocides, Section 9.1.4 of the current Guidance on BPR Volume IV - Part A specifies that *"[t]he experimental determination may not need to be carried out if it can be demonstrated on the basis of physico-chemical properties (e.g. log K_{ow} <3) or other evidence that the substance has a low potential for bioconcentration. All critical aspects of bioaccumulation such as ionic speciation, surface activity and metabolic transformation rates must be considered before experimental determination is considered unnecessary*.".

13. Fish toxicity

13.1 Information requirements

13.1.1 Under the REACH Regulation

As described in REACH Annex VI, all available existing information should be collected and considered in the hazard assessment, regardless of whether testing for a given endpoint is required or not at a specific tonnage level.

Minimum information requirements are set out in Annexes VII- X. If information required in Annexes VII-X is not available, testing is required unless a modification according to general rules described in Annex XI is possible.

Information on the short-term toxicity to fish is required when a substance is registered in quantities higher than 10 tonnes per year. However, a short-term fish toxicity test does not need to be performed if results from a long-term fish toxicity study are already available. In addition, a short-term fish toxicity study is not needed if there are mitigating factors indicating that aquatic toxicity is 'unlikely to occur', for example, when the substance is highly insoluble in water or when the substance is unlikely to cross biological membranes. However, if the substance is poorly water soluble, a long-term aquatic toxicity study in fish (Annex IX, Section 9.1.6) must be considered instead of the short-term test.

Information on long-term toxicity to fish is required when a substance is registered in quantities higher than 100 tonnes per year. However, long-term toxicity testing in fish (Annex IX, 9.1.6) can be omitted if the chemical safety assessment according to Annex I does not indicate the need to further investigate the effects on aquatic organisms and sufficient scientific justification is provided.

13.1.2 For biocide active substances under the BPR

The data elements belonging to the core dataset are considered as the basic data that should, in principle, be provided for all active substances. However, in some cases, due to the physical or chemical properties of the substance it may be impossible or unnecessary to provide specific data elements belonging to the core dataset.

Short-term toxicity testing on fish is part of the core dataset and must always be provided. It is nevertheless specified that when short-term fish toxicity data is required, the threshold approach (tiered strategy) should be applied and that the study does not need to be conducted if a valid long-term aquatic toxicity study on fish is available.

Long-term toxicity testing on fish is an additional data requirement (part of the additional dataset) only for certain product-types (where, for example, continuous release to the aquatic compartment occurs).

With regard to the additional dataset, the data elements to be provided for a specific active substance should be determined by taking into account the physical and chemical properties of the substance, existing data, information which is part of the core dataset and the types of products in which the active substance will be used and the exposure patterns related to these uses.

For some of the information requirements, it may be possible to satisfy the requirements based on available information of the properties of the active substances contained in the product. Where valid data on the components are not available or where synergistic effects may be expected, then testing of components and/or the biocidal product itself may be necessary.

13.2 CLP criteria

The CLP Regulation is concerned with determining hazard, i.e. the presence or absence of a toxic effect at specific hazard endpoints and whether toxicity occurs at a concentration within predetermined criteria, for environmental hazards.

As such, test guidelines used for hazard identification under the CLP Regulation must satisfactorily show toxicity at hazard endpoints and the concentration at which the effect occurs. Many other types of data can be submitted under the CLP Regulation, but will only be considered under a WoE approach (see Section 4.1 of the Guidance on the application of the CLP criteria).

13.3 Relevant test methods

Under the REACH Regulation, the following test methods can be used (see also Section R.7.8.3 and Appendix R.7.8-2 of the Guidance on IR&CSA – Chapter R.7b):

- Fish, Acute Toxicity Test (EU C.1/OECD TG 203) is the preferred test to cover the standard information requirement of Annex VIII, Section 9.1.3, i.e. a short-term study.
- Fish short-term reproduction assay (OECD TG 229) is a 21-day screening study for substances that affect reproduction through various mechanisms, including endocrine modalities (i.e. oestrogenic and androgenic activity and aromatase inhibition). This study is not a standard information requirement, but it is accepted as a screening study to detect potential endocrine disruptors (See Section B.11 on endocrine-disrupting properties and Section 11 of Appendix 3 to this document).
- 21-day fish assay (OECD TG 230), a short-term screening for oestrogenic and androgenic activity and aromatase inhibition. This is not a standard information requirement, but it is accepted as a screening study to detect potential endocrine

disruptors (See Section B.11 on endocrine-disrupting properties and Section 11 of Appendix 3 to this document).
- The fish early-life stage toxicity test (OECD TG 210), fish short-term toxicity test on embryo and sac-fry stages (EU C.15/OECD TG 212) and fish juvenile growth test (EU C.14/OECD TG 215) are the preferred tests to cover the standard information requirement of Annex IX, Section 9.1.6, i.e. a fish long-term study.

Note: Regarding the long-term toxicity testing on fish according to Annex IX, Section 9.1.6.1, ECHA considers that the fish early-life stage toxicity test according to OECD TG 210 is a sensitive test as it covers several life stages of the fish from the newly fertilised egg, through its hatch to the early stages of growth. Hence, a preference is generally given for using this test method when long-term fish toxicity testing is warranted (see Figure R.7.8-4 of the Guidance on IR&CSA - Chapter R.7b). ECHA considers the fish early-life stage toxicity test using the OECD TG 210 as the most appropriate and suitable.

- The fish sexual development test (FSDT) (OECD TG 234) is another fish early-life stage test. It is an enhancement of OECD TG 210, where the exposure is continued until the fish are sexually differentiated, and sex hormone relevant endocrine-sensitive endpoints are investigated. It can be requested, if the substance or its metabolites are suspected of having endocrine-disrupting properties.

- For difficult to test substances, OECD GD 23 on Aquatic Toxicity Testing of Difficult Substances and Mixtures [239] and Table R.7.8-3 of the Guidance on IR&CSA – Chapter R.7b, summarising aquatic toxicity testing of difficult substances for choosing the design of the requested ecotoxicity tests and for calculating and expressing the result of the tests, should be consulted.

Regarding biocides, for the short-term fish toxicity data, one species should be tested as part of the core dataset, preferably a fresh water species or, if different aquatic environments are exposed, two species may be required. The recommended guidelines are OECD TG 203 or the US EPA guideline OPPTS 850.1075 (Fish Acute Toxicity Test, Freshwater and Marine). Short-term study does not need to be conducted if a valid long-term aquatic toxicity study on fish is available. Also, the threshold approach (tiered strategy) according to the OECD GD 126 [228] must be considered.

In addition to the tests mentioned for REACH, a fish full life cycle test (FFLCT) may be necessary for biocides, if results from other long-term studies with fish indicate a concern (see also Section 9.10 of the Guidance on BPR Volume IV - Part A, Chapter II "Identification of endocrine activity").

OECD TG 240 "*Medaka Extended One Generation Reproduction Test (MEOGRT)*" can be used for a fish full life cycle test. Two reviews of existing testing approaches and protocols under development are also available: the OECD series on testing and assessment No. 95 "*Detailed Review Paper on Fish Life-cycle tests*" [240] and No. 171 "*Fish Toxicity Testing Framework*" [241], including the one-generation fish full life cycle test likely to be sufficient to satisfy regulatory requirements.

Further toxicity studies on aquatic organisms (additional dataset) should be conducted if a need is indicated by other elements of the assessment (e.g. long-term exposure expected or clear exposure to marine or brackish environment). There are also product-type specific requirements, which are specified in Annex V to the Guidance on BPR Volume IV - Part A.

For CLP, the tests accepted for hazard classification are the same tests as those listed for REACH above.

13.3 Specific adaptation rules

Short-term aquatic toxicity to fish

The REACH Regulation (Annex VIII, 9.1.3) describes the circumstances under which short-term toxicity testing on fish is not necessary and the information requirement may therefore be adapted with justifications.

If a long-term toxicity study is already available for fish, a short-term study is not needed.

In addition, a short-term study is not needed if there are mitigating factors indicating that aquatic toxicity is unlikely to occur, for example, when the substance is highly insoluble in water or when the substance is likely not to cross biological membranes. However, if the substance is poorly water soluble, a long-term aquatic toxicity study on fish (REACH Annex IX, Section 9.1.6) must be considered instead of the short-term test.

In relation to biocides (BPR Annex II, Section 9.1.1), the threshold approach (tiered strategy) for acute fish toxicity testing [228] should be applied or the requirement for this study should be adapted if a valid long-term aquatic toxicity study on fish is available.

Long-term aquatic toxicity to fish

For information requirements on long-term toxicity testing on fish (Annex IX, 9.1.6), long-term toxicity testing can be omitted if the chemical safety assessment according to Annex I does not indicate the need to investigate the effects on aquatic organisms further.

A chemical safety assessment should include persistent, bioaccumulative and toxic (PBT)/very persistent, very bioaccumulative (vPvB) assessment and an environmental hazard assessment, including classification and labelling in accordance with the CLP Regulation and determination of the predicted no-effect concentration (PNEC) used in risk assessment. Therefore, to omit long-term toxicity testing on fish, PBT assessment, classification and labelling, and risk assessment need to be considered.

For risk assessment purposes, a risk from CSA also includes criteria that are not related to risk characterisation ratios alone. A risk from CSA is indicated when predicted environmental concentrations (PECs) exceed the PNEC, but also when the log K_{ow} of a substance exceeds 3 (or BCF >100) and a PEC_{local} or $PEC_{regional}$ >1/100th of the water solubility.

In addition, results from a qualitative assessment may show a risk e.g. information on a specific mode of action (MoA) and unexpected sensitivity of a group of organisms to the substance under investigation is identified.

Whichever argument is used for not performing a test it should be accompanied by a clear, scientific explanation in the registration dossier.

EUROPEAN CHEMICALS AGENCY
ANNANKATU 18, P.O. BOX 400,
FI-00121 HELSINKI, FINLAND
ECHA.EUROPA.EU