National Institute for Public Health
and the Environment
Ministry of Health, Welfare and Sport

Innovating the integration of ecological and human health Risk Assessment: Connecting concepts and cases (IRAC) - Identification phase, parallels and integration

RIVM letter report 2020-0010
J.A. de Knecht | M. Marinković | Z. Dang

National Institute for Public Health
and the Environment
Ministry of Health, Welfare and Sport

Innovating the integration of ecological and human health Risk Assessment: Connecting concepts and cases (IRAC) - Identification phase, parallels and integration

RIVM letter report 2020-0010
J.A. de Knecht | M. Marinković | Z. Dang

Colophon

RIVM attaches a great deal of importance to the accessibility of its products. However, it is at present not yet possible to provide this document in a completely accessible form. If a part is not accessible, it is mentioned as such. Also see www.rivm.nl/toegankelijkheid.

DOI 10.21945/RIVM-2020-0010

J. de Knecht (author), RIVM
M Marinković (author), RIVM
Z. Dang (author), RIVM

Contact:
Joop de Knecht
Centre for Safety of Substances and Products
Joop.de.knecht@rivm.nl

This investigation was performed by order, and for the account of the Director General of the RIVM, within the framework of the Strategic Program RIVM

Published by:
National Institute for Public Health and the Environment, RIVM
P.O. Box 1 | 3720 BA Bilthoven
The Netherlands
www.rivm.nl/en

Synopsis

Innovating the integration of ecological and human health Risk Assessment: Connecting concepts and cases (IRAC) - Identification phase, parallels and integration

There are methods to assess risks to humans and methods to assess risks to the environment. Although both involve risk assessment, the test methods and objectives used are different. Nevertheless, by integrating the two types of risk assessment, risk analysis could be strengthened with the possibility of reducing animal testing, prioritizing substances better in order to make a more efficient policy decision.

RIVM has compared the assessment methods to determine the differences and similarities and describes the possibilities for achieving more integration. Based on this, it can be concluded that the data and models used to predict the emission and distribution of substances are reasonably comparable and integrated for both risk assessment frameworks. As experimental animals may be exposed differently, extrapolation between species should use more internal rather than external concentrations. For this, use could be made of models that simulate the uptake, distribution, metabolism and elimination of a substance after ingestion by a species. These models, together with models that determine the dose response, can also help to translate the results of alternative (in vitro) test methods into the in vivo situation. Relatively little attention has been paid to the similarities in the processes that underlie the toxicity within the different species, mainly because the knowledge about the mechanism of action was lacking. Due to the recent advances in molecular and biological science, this knowledge is increasing. If there is sufficient evidence that the toxicity in humans and other animal species follow the same adverse outcome pathway, this can be used to extrapolate effects from one species to another and to make better use of alternative testing methods in the risk assessment.

Keywords: integrated risk assessment, Adverse Outcome Pathways, environmental chemicals, (eco)toxicology, alternatives to animal testing, toxicokinetics, exposure

Publiekssamenvatting

Innovatie van de integratie van ecologische en humane risicobeoordeling: verbindende concepten en cases (IRAC) - identificatiefase, parallellen en integratie

Er bestaan aparte methoden om risico's van stoffen voor mensen en voor het milieu te beoordelen. Hoewel beide methoden risico's beoordelen, gebruiken ze andere testmethoden en verschillen de doelen (denk aan effecten op organen bij mensen versus effecten op groepen dieren in het milieu). Toch kan de kwaliteit van een risicoanalyse van een stof beter worden als de resultaten van beide typen risicobeoordelingen worden samengevoegd. Dan kunnen de methoden nog beter aangeven voor welke stoffen de overheid met voorrang beleid moet maken. Ook zijn er mogelijk ook minder dierproeven nodig.

Het RIVM heeft daarom de verschillen en overeenkomsten van de risicobeoordelingen voor mens en milieu in kaart gebracht. Hieruit blijkt dat ze vergelijkbare gegevens en modellen gebruiken om te voorspellen hoeveel van een stof in het milieu terechtkomt en aan welke hoeveelheid stoffen mens en dier worden blootgesteld. De hoeveelheden die in verschillende diersoorten terechtkomen zijn echter niet gelijk. Als we diersoorten met elkaar willen vergelijken, is het noodzakelijk om de gehaltes te kennen die in het dier of de mens zitten. Om deze gehalten in een proefdier te bepalen, kunnen modellen worden gebruikt die nabootsen hoe het proefdier stoffen opneemt, verdeelt, verteert en uitscheidt.
Lange tijd was er in de wetenschap weinig bekend over de manier waarop stoffen schadelijke effecten veroorzaken in proefdieren en of dat bij verschillende diersoorten op dezelfde manier gebeurt (werkingsmechanisme). De kennis hierover neemt de laatste jaren sterk toe. Wanneer de werking voldoende overeenkomt, kunnen testresultaten van de ene naar de andere soort worden doorvertaald. De resultaten zouden dan kunnen worden gecombineerd met modellen die aangeven hoe een dier op een interne dosis van een stof reageert. Dit kan ook helpen de resultaten van andere (reageerbuis) testmethoden te vertalen naar de reactie in cellen en in het hele organisme.

Kernwoorden: geïntegreerde risicobeoordeling, Adverse Outcome Pathways, chemicaliën, (eco)toxicologie, alternatieve voor dierproeven, toxicokinetiek, blootstelling, werkingsmechanisme

Contents

Summary

Human risk assessment (HRA) and environmental risk assessment (ERA) have developed separately, and while they share the same risk assessment paradigm, they differ in applied approaches and their protection goals. Harmonization of both types of risk assessment with the eventual goal to achieve an integrated risk assessment is generally considered the way forward. Not only would this decrease the pressure on available resources and reduce the number of animals tested, but it would also support a more mechanistically based risk assessment. This could reduce the uncertainties associated with the current animal-based approaches, and would facilitate extrapolation among species. The project named IRAC (full name: Innovating the integration of ecological and human health Risk Assessment: Connecting concepts and cases) aimed to advance the integration of the HRA and ERA by setting up a framework based on innovative concepts.

In this report, a first step was made by aligning both types of risk assessment to determine the commonalities and complementary elements as well as the gaps and redundancies, with the aim to identify opportunities for integration that was worked out in later stages of the IRAC project as case studies. A prerequisite for comparison is to verify that essential terms and/or concepts have equal meanings in both types of risk assessment. Chapter 3 shows that while there are slight differences in substance definition and substance identification between legislative frameworks, within a framework there are no differences between the HRA and ERA. The same applies for the definition of physicochemical and solvation properties of a substance. As discussed in Chapter 4, for such basic knowledge past cross-fertilization has resulted in standardized test guidelines and QSAR models, and further integration is not expected to result in substantial improvements or efficiency increases. The usage of physicochemical and solvation data as input in *in silico* models to estimate various processes, such as, emission and distribution of substances, exposure levels, toxicokinetic and toxicodynamic processes, and exerted toxicity, is tailored to the respective needs of the type of risk assessment and the legislative framework. Integration is often though achievable, as demonstrated by the EU FP7 project HEROIC and the paper on aggregated exposure pathways (AEPs) that both advanced on the integration of human and environmental exposure assessment. As concluded in Chapter 5 these achievements do not warrant further efforts to be taken in the current project to further integrate human and environmental exposure assessment.

Opportunities identified for integration of the HRA and ERA have been identified linked to toxicology and the underlying toxicokinetics and toxicodynamics processes. While the HRA aims to protect individuals and mechanistic knowledge is generated, in the ERA protection is aimed at the level of the individual and considerable effort is pointed towards mechanistic insight.

A more challenging area of potential cross-fertilization between HRA and ERA experts is the estimation of internal doses and subsequent concentrations at target sites. Considering that toxicokinetic studies are often not a standard requirement in HRA and are not frequently performed in the ERA, it is considered very informative when internal concentrations from one species could be extrapolated to another species, within or across taxa. This could support the estimation of potential hazards in untested species, based on the assumption that the internal concentrations are indicative for acute toxicity and the toxicity at the target site. A concrete example worth exploring would be the relationship between a substance bioaccumulation potential in fish, its accumulation in rodents, and its half-life in human blood.

How these physicochemical and solvation data are subsequently used as input in *in silico* models to estimate exposure, toxicokinetics & toxicodynamics processes, and toxicity appears to be tailored to the respective type of risk assessment. That said cross-fertilization takes place and efforts are made to translate in silico approaches and methods from HRA to ERA and vice versa with respect to exposure assessment, toxicokinetics & dynamics, and toxicology.

1 Introduction

In human and environmental risk assessment, hazard-driven and animal-based approaches are applied that are demanding in terms of time, money and animal use. The current approaches suffice for risk assessment purposes, but yield relatively little information on the mechanisms underlying toxicity. This hampers extrapolation among species, and results in high uncertainties about actual risks. When also taken into account that there is an ever-increasing number of chemicals that enter the market while resources remain at best at the same level, it is rather evident that there is a need for increased efficiency in risk assessment.

Historically, the procedures for human risk assessment (HRA) and for environmental risk assessment (ERA) have developed separately with different terminologies and separate databases. On a more fundamental level, there are also differences between the two types of risk assessment. The most prominent one being that the HRA focuses on protecting individuals, whereas the ERA is more concerned about populations of organisms, communities of selected organisms and the entire ecosystem. Consequently, the HRA applies a considerably more conservative approach, which reflects a difference in risk perception; the death of a human is not considered acceptable, the exception being that for mutagenic carcinogens one in a million additional cancer cases is considered acceptable, while in the ERA it is accepted that not all individuals can be protected. To achieve these protection goals there is also a difference in how toxicity data is used. In the HRA several species are tested to assess potential human health effects whereas in the ERA only a few indicator species are used to evaluate potential effects on many different species and even the ecosystem as whole. The HRA and ERA also differ regarding the level at which toxicity is studied. HRA mostly evaluates effects at tissue, cellular and molecular levels, thus considering toxicodynamics and toxicokinetics, whereas the focus of the ERA is on less specific toxicity endpoints that are observed at higher levels of biological organization such as population growth and survival. On the exposure side the two types of risk assessments also differ, with the HRA considering all relevant routes of exposure whereas in practice the ERA tends to focus on single routes of exposure even where the potential for other routes is clear (e.g. oral vs. dermal exposure of birds to pesticides).

Besides the above discussed differences, there are also aspects that both types of risk assessments have in common. Not only are they linked, as humans can be seen as part of the ecosystem, but biochemical pathways tend, at least to some extent, to be conserved across species. Furthermore, the same exposure routes may apply for humans and other organisms. Perhaps the most important aspect that the HRA and ERA share, is that they follow the same risk assessment paradigm, i.e. risk is estimated by comparing hazard data with estimates of exposure. Overall, it is believed that a harmonization of both types of risk assessment is feasible and that an integrated assessment could be achieved.

In a joint workshop organized by the UNEP/ILO/WHO International Programme on Chemical Safety (IPCS), in collaboration with the US-EPA, the European Commission, the OECD and other (inter)national organizations the following major advantages of integrated risk assessment were identified (Suter, Vermeire et al. 2001, Suter, Vermeire et al. 2003):

1. Coherent expression of assessment results, which facilitates the decision-making process.
2. It is easier to account for interdependencies between human health and ecosystems in an integrated approach.
3. Integrated assessment may enhance the use of nonhuman organisms as sentinels to identify potential sources of human hazards (Burger and Gochfeld 2001).
4. The scientific quality of the assessments is expected to improve through the sharing of information and techniques between scientists of different disciplines.
5. The efficiency of the assessment will improve because data gathering and model applications will be harmonized.

More recently, Wilks, Roth et al. (2015) concluded that an integrated approach offers the perspective of harmonized risk assessment models and methodologies and a better crosstalk between HRA and ERA. The mutual exploitation of all existing data across the two disciplines could inform the risk analysis process better on the overall risks to human health and the environment for a more comprehensive and efficient health- and environmental-based decision making. This would subsequently allow for a better prioritizing and ranking of all potential risks and could contribute to more adequate risk management decisions.

A central feature of IRAC (full name: Innovating the integration of ecological and human health Risk Assessment: Connecting concepts and cases) is the bringing together of independent sources of toxicological and ecotoxicological data that are usually kept separate. This includes standard guideline sources, but also information from non-guideline sources, non-testing and modelling data, and data obtained for non-model species. Such integration enables a more comprehensive, efficient and informative risk assessment and provides the opportunity to exploit the best aspects of both worlds. The approach followed in IRAC is to develop an efficient, 3R-compliant and flexible risk assessment framework integrating ecological and human health, while introducing innovative concepts such as adverse outcome pathway (AOP) concepts (Ankley, Bennett et al. 2010) and the aggregated exposure pathway (AEP) concepts (Teeguarden, Tan et al. 2016).

1.1 Objective

In this report, current regulatory approaches in environmental and human health risk assessment of chemicals have been compared. Both approaches have been aligned in parallel, and commonalities and complementary elements have been mapped. Figure 1 visualizes how the comparison of both lines of risk assessment have been investigated. When applicable, gaps and redundancies in both approaches have been identified. Based on the comparison, opportunities for integration of human health and environmental risk assessment have been identified,

leading to possibilities to enhance the efficiency of risk assessment procedure for man and the environment.

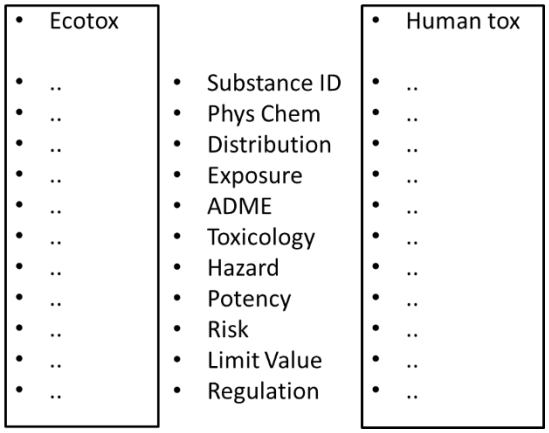


Figure 1 Template for mapping parallels between human (HRA) and environmental (ERA) risk assessment components

An early integrated framework for human and environmental risk assessment was designed by IPCS (http://www.who.int/ipcs/methods/risk_assessment). This framework integrates components such as (1) exposure and effects; (2) *in silico*, *in vitro*, *in vivo* or monitoring data; (3) multiple chemicals, multiple species/target organisms, multiple toxicological endpoints, multiple exposure routes; (4) spatial and temporal scales; (5) a product's life cycle; (6) or socio-economic aspects (Suter, Vermeire et al. 2003). Like most other risk assessment frameworks, it consists of four basic elements or steps: problem definition, characterization of exposure, characterization of effects and characterization of risk (Figure 2). These steps would include:

- The development of an integrated analysis plan during the problem definition phase, which defines the sources, stressors, endpoints, scale levels, methods and conceptual models included in the assessment.
- The integrated characterization of sources, emissions, fate and exposure for human and ecological receptors. Some pathways are unique to humans (e.g., contamination of food packaging) or ecological receptors (e.g., drinking from waste sumps), but there is much commonality in the characterization of exposure.
- The characterization of (common) modes of action of stressors in organisms can improve dose response modelling, the use of biomarkers and indicators, and the extrapolation between observed and estimated endpoints (characterization of effects).
- The use of similar procedures and methods to determine causation, combine lines of evidence, quantify uncertainty and present the results (characterization of risks).

Perceived benefits were consistent expressions of assessment results, incorporation of the interdependence of humans and the environment, use of sentinel organisms, and improvement of efficiency and quality of assessments.

The advantage of the framework to toxicologists was, amongst others, the opportunity to use understanding of toxicokinetics and toxicodynamics to perform the integrated assessment of all exposed species. The IPCS proposal had limited practical follow-up. Though considered conceptually useful, implementation by stakeholders remained limited, reflecting that policy problems were still considered as either human health or environmental risk assessment. The implementation barrier was not considered regulatory in nature, as in most nations the laws that control pollution assure the quality of air, land and water for protecting both human health and the environment (Vermeire, Munns et al. 2007).

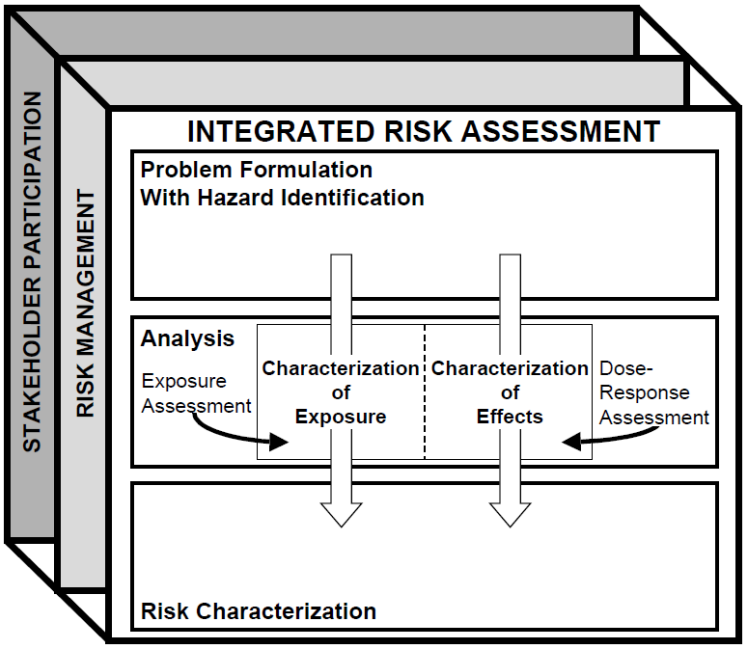


Figure 2 A framework for integrated human health and ecological risk assessment (obtained from (Suter, Vermeire et al. 2001, Suter, Vermeire et al. 2003).

Risk assessors, risk managers, and stakeholders perform parallel activities which may interact at various stages.

To assess exposure of and toxicity to humans and biota in similar terms a common modelling framework is needed based on a (mechanistic) understanding of the exposure pathways and the toxicokinetics and toxicodynamics of chemicals. To achieve this goal the concept of aggregated exposure pathway (AEP) proposed by Teeguarden, Tan et al. (2016) and the concept of adverse outcome pathway (AOP) proposed by Ankley, Bennett et al. (2010), could be helpful frameworks. The principle components of an AEP cover all necessary levels of ecological, biological, and physical organization from sources to target site exposure (see paragraph 4.3). The AOP concept describes a sequential chain of causally linked events at different levels of biological organization from the molecular initiating event leading to an adverse human health or ecotoxicological effect. Together, the two frameworks form and inform a decision-making framework allowing for risk-based, hazard-based, or exposure based decision making (Teeguarden, Tan et al. 2016).

2 Substance Identification

2.1 Substance definition

Prior to the identification of a substance it must be clear what is meant by the term substance. The meaning can differ slightly between legislative frameworks, as is shown below, but within a framework the term is interpreted equally for the human health and the environmental part of the risk assessment of chemicals, if both are to be conducted.

2.1.1 Substance definition under the REACH and CLP regulations

Under the REACH (Registration, Evaluation, Authorization and Restriction of Chemicals) regulation (EC No. 1907/2006) a substance is defined under article 3(1) as a chemical element and its compounds in the natural state or obtained by any manufacturing process, including any additive necessary to preserve its stability and any impurity deriving from the process used, but excluding any solvent which may be separated without affecting the stability of the substance or changing its composition. The same definition is used in the CLP (classification, labelling and packaging of substances and mixtures) regulation (EC No. 1272/2008) under article 2(7). Thus, under REACH and CLP the definition of a substance goes beyond a pure chemical compound composed of a single molecule. The term covers both substances obtained by a manufacturing process and substances in their natural state, both of which can consist of several constituents within the substance. These constituents can be one or more main constituents, but also impurities, i.e. unintentional constituents coming from the manufacturing process or from the starting material(s), or additives, i.e. constituents which are intentionally added to stabilise the substance and only for this purpose. Provided that the composition of a substance can be quantitatively and qualitatively defined, it can be considered a mono-constituent substance, i.e. one constituent is present at a concentration of at least 80% (w/w) and the impurities make up no more than 20% (w/w), or a multi-constituent substance, i.e. more than one main constituent is present in a concentration between 10% and 80% (w/w). Then there is a category of substances whose composition cannot be sufficiently identified as it contains too many constituents, the composition is to a significant extent unknown, or the variability of composition is large or unpredictable. Such substances are considered UVCB (Unknown or Variable composition, Complex reaction products or Biological materials) substances (ECHA 2014).

2.1.2 Substance definition under the Biocidal Product regulation

Within the biocides framework under the Biocidal Product (BP) Regulation (528/2012/EC) the definition of a substance is equal to that applied in the REACH regulation. In addition, the term active substance is used, which is defined under article 3(c) of the BP regulation as a substance or micro-organism that has an action on or against harmful organisms.

2.1.3 *Substance definition under the Plant Protection Products regulation*

Under the Plant Protection Products (PPP) regulation (1107/2009/EC) the definitions differ slightly from those used in REACH. Substances are defined under article 3 as chemical elements and their compounds, as they occur naturally or by manufacture, including any impurity inevitably resulting from the manufacturing process, while impurities are defined as any components other than the pure active substance and/or variant which is present in the technical material (including components originating from the manufacturing process or from degradation during storage). In analogy to the BP regulation, the PPP regulation identifies active substances as substances, including micro-organisms having general or specific action against harmful organisms or on plants, parts of plants or plant products. The PPP regulation defines several other types of substances, i.e. safeners, synergists, co-formulants, and adjuvants, whose definitions will not be given here, but can be found in Article 2 of the PPP regulation. Products containing one or more active substances with the aim to protect plants or plant products against harmful organisms or preventing the action of such organisms (before or after harvest), to influence the life processes of plants (excl. nutrients), to preserve plant products, to destroy undesired plants or parts of plants and/or to check or prevent undesired growth of plants, are known as PPP or 'pesticides'.

2.1.4 *Substance definition under the Veterinary and Human Medicinal Products directives*

The Veterinary Medicinal Products (VMP) Directive (2001/82/EC) and the Medicinal Products for Human Use (HMP) Directive (2001/83/EC) define substances as any matter irrespective of origin which may be: (1) human, e.g. human blood and human blood products; (2) animal, e.g. micro-organisms, whole animals, parts of organs, animal secretions, toxins, extracts, blood products; (3) vegetable, e.g. micro-organisms, plants, parts of plants, vegetable secretions, extracts; (4) chemical, e.g. elements, naturally occurring chemical materials and chemical products obtained by chemical change or synthesis. Both directives refer to active substances and excipients. Active substances can be classified as new or existing, depending if they are being used for the first time in a medicinal product either for human of veterinary use. Furthermore, existing active substances can be described in the European pharmacopoeia (Ph.Eur) or the pharmacopoeia of an EU member (EMEA 2004). Excipients are constituents of a pharmaceutical form apart from the active substance, and can include fillers, disintegrants, lubricants, colouring matters, antioxidants, preservatives, adjuvants, stabilisers, thickeners, emulsifiers, solubilisers, permeation enhancers, flavouring and aromatic substances etc., as well as the constituents of the outer covering of the medicinal products, e.g. gelatine capsules (EMEA 2007).

2.1.5 *Substance definition under the Cosmetic Products Regulation*

Under Article 2 of the Cosmetic Products Regulation (EC No. 1223/2009) the following definition is used, substance means a chemical element and its compounds in the natural state or obtained by any manufacturing process, including any additive necessary to preserve its stability and any impurity deriving from the process used but excluding any solvent which may be separated without affecting the stability of the

substance or changing its composition. Several substances are specifically defined, i.e. preservatives, colorants and UV-filters.

2.1.6 Substance definition under the Food and feed additives Regulations

Under the Food Additives Regulation (EC No. 1333/2008) and Feed Additives Regulation (EC No. 1831/2003) the term substance is not defined, but more specifically the terms food and feed additives, respectively. Food additives are substances that are not normally consumed as food itself but are added to food intentionally for a technological purpose described in the Food Additives Regulation, such as the preservation of food.

2.1.7 Water Framework Directive

Water Framework Directive (2000/60/EC) only defines hazardous and priority substances, and pollutants.

2.1.8 Groundwater Directive

Directive on the protection of groundwater against pollution and deterioration (2000/60/EC) does not define what a substance is, nor specific groups of substances.

2.2 Substance identification

As the term substance does not refer to a pure chemical compound composed of a single molecule, unambiguous and clear substance identification is an essential preliminary step in risk assessment. While the different frameworks discussed above might differ slightly in their requirements, the parameters needed to describe the identity of substances are pretty much universal. Generally, a substance is identified by its chemical composition, the chemical identity and the content of each constituent in the substance. Under the REACH regulation the following substance identification parameters are specified in Annex VI:

1. Name or other identifiers of a substance, e.g. IUPAC name, other chemical name(s), usual name, trade name, abbreviation, CAS name, CAS number, EC number, other identity code (REACH, Annex VI, section 2.1);
2. Information related to molecular and structural formula of a substance, e.g. molecular and structural formula, SMILES notation, information on optical activity and typical ratio of (stereo) isomers, molecular weight or molecular weight range (REACH, Annex VI, section 2.2);
3. The chemical composition of a substance, e.g. degree of purity (%); percentage and nature of impurities (isomers and by-products) and additives; spectral data (ultra-violet, infra-red, nuclear magnetic resonance or mass spectrum (other spectroscopic methods are possible, e.g. atomic absorption spectroscopy)) to confirm the structure of the substance; high-pressure liquid chromatogram and/or gas chromatogram (other separation techniques are also possible) to confirm the composition of the substance; description of analytical methods or references (REACH, Annex VI, section 2.3).

Although for most mono- or multi-constituent substances the above parameters will allow a complete identification, for some substances a straightforward identification may not be possible as they are defined by more than the chemical composition, and other or additional substance identification information is needed. This can be information on crystallomorphology and (geological) mineral composition, as demonstrated by graphite and diamond (ECHA 2014), but also information on particle size and particle size distribution, particle shape/aspect ratio, aggregation/agglomeration states, porosity and specific surface area of particles, which can serve as potential identifiers for nanomaterials (JRC 2011, OECD 2016). From the advisory report that was drafted by the Joint Research Centre based on work carried out by an expert group comprised by participants from Member State Competent Authorities, industry, NGOs and ECHA, it is clear that additional information is needed for identification and naming of nanomaterials, but also that no consensus could be reached on the type of information (JRC 2011). More details on the identifiers can be found in the section on physicalchemical properties of nanomaterials (see section 3.1.11).

For UVCB substances, the basic rule is that substances should be defined as much as possible by the chemical composition and the identification of the constituents. The description of the composition can often be given in a more generic way, but all known constituents and all constituents present at concentrations ≥ 10% should be specified by an IUPAC name and preferably a CAS number, and the typical concentrations and concentrations ranges of the known constituents should be given as well. Only if this is not technically feasible other identifiers should be used for the various types of UVCB substances, where the substance shall in general be identified by its chemical name, its origin or source (e.g. biological, chemical or mineral source) and the most relevant steps taken during processing (e.g. synthesis, refinement). Other substance properties can also be important identifiers, either as relevant generic identifiers (e.g. boiling point) or as crucial identifiers for specific groups of substances (e.g. catalytic activity for enzymes). For more information, the reader is referred to the guidance document for identification and naming of substances under REACH and CLP (version 1.3; Feb. 2014) (ECHA 2014).

For active substances under the BP and PPP regulations, basically the same applies as under REACH, with the addition that the analytical profile needs to be determined in five representative batches for BP, and in five batches for each manufacturing location for PPP. Furthermore, the PPP regulation requires that inactive isomers are specified and identified.

In the VMP and HMP directives it is stated that the monographs of the European Pharmacopoeia (Ph.Eur) shall be applicable to all substances, i.e. active substances and excipients, appearing in it. For substances not appearing in the European Pharmacopoeia (Ph.Eur) or the pharmacopoeia of an EU member, the description should be in the form of a monograph according to Ph. Eur, which includes the following information: International Nonproprietary Name (INN) established by the World Health Organization; the definition of the substance, i.e.

information related to molecular and structural formula of a substance, and the method of synthesis; the chemical name; purity of the substance; description of the analytical methods used for identification, i.e. spectrophotometry (ultra-violet, visible, infra-red), liquid chromatography and gas chromatography, thin layer chromatography, specific optical rotation, melting point, freezing point and boiling point, and chemical reactions.

Under the cosmetics regulation it is specified that the qualitative and quantitative composition of the cosmetic product, including chemical identity of the substances (incl. chemical name, INCI, CAS, EINECS/ELINCS, where possible) and their intended function should be reported. It is however, not detailed how the chemical composition should be determined.

3 Physicochemical and solvation properties

Before the hazard and risk potential of a substance to human health and the environment can be evaluated, a good understanding of its physicochemical and solvation properties is needed. Such information is required to assign physical hazard classes, design a correct setup for experimental fate and effect studies, and as input for *in silico* approaches, e.g. QSARs, distribution, exposure and/or metabolism models. While there is overlap in how data on physicochemical and solvation properties is used in human health and environmental risk assessment, there are also differences. Both will be discussed in this chapter, but first a brief description will be given of the most critical physicochemical and solvation properties, and the processes they affect.

3.1 Relevant physicochemical and solvation properties

There are differences between frameworks, but a general basic knowledge on physicochemical properties is required by all. Under the REACH regulation the following physicochemical information is considered relevant, and has to be included in the technical dossier (ECHA 2016):

- data on intrinsic properties of the substance (e. g. melting /freezing point, boiling point, vapour pressure, density);
- data necessary to assess the physical hazards of a substance (e. g. flammability, explosiveness, oxidising properties);
- supplementary data for hazard assessment and health and environmental classification (e. g. viscosity, *n*-octanol/water partition coefficient).

Information on solvation properties is required under REACH, even though it is grouped under supplementary data and not specified as such. Solvation properties describe a chemical's interaction with different phases and its partitioning between phases, and can be divided into three main types: phase partitioning (e.g. *n*-octanol/water partition coefficient), solubility (e.g. water solubility), and colligative properties (National Research Council 2014). The first two types are evident and need no further explanation. Colligative properties on the other hand are less clear, as they are properties of solutions that are not dependent on the substance, but instead on the ratio of the number of solute particles to the number of solvent molecules in a solution, examples are lowering or elevating of melting, freezing and/or boiling points (National Research Council 2014). Colligative properties are not always investigated and reported for substances.

Generally spoken when determining physicochemical and solvation properties, care should be taken with regard to impurities, as these can influence the properties significantly. In the sections below, the provided information is largely based on the REACH regulation. Where needed it is extended, and the respective sources are specified. This section does not intend to cover everything stated under REACH Guidance document R7a (ECHA 2016), but only focusses on the most relevant physicochemical and solvation properties with regard to human health

and environmental risk assessment. It should also be noted that there can be interdependence between physicochemical properties, where appropriate, this is indicated in the relevant sections.

3.1.1 General substance characteristics: physical state, appearance and odour

The most basic physicochemical properties are the physical state, appearance & odour of a substance. The physical state of a substance, i.e. being a solid, liquid or gas, depends greatly on its melting and boiling points, as is discussed in the sections below (sections 3.1.2 and 3.1.3). Substance can be present in the same physical state and can still differ in their appearance, i.e. they can consist of smaller particles, i.e. flakes or powder, both of which can consist of different particle sizes and shapes, or differ in crystallinity and/or allotropy. This is especially relevant for nanomaterials that have been engineered to consist of very small particles often with specific properties. Therefore, while granulometry is not a specific physicochemical property of a substance, it is discussed below (section 3.1.11). Other general characteristics of substance are its colour and/or odour.

3.1.2 Melting point

The melting point is defined as the temperature at which the phase transition from the solid to the liquid state occurs at atmospheric pressure and this temperature ideally corresponds to the freezing temperature. Methods suitable to determine the melting point experimentally are described in OECD TG 102 (OECD 1995). Under REACH, the melting point does not need to be determined experimentally if it is below -20°C. The lower limit of -20°C needs to be determined experimentally though, unless a melting point below -50°C is estimated by (Q)SAR. Application of (Q)SARs for the determination of the melting point is discouraged, due to low accuracy (± 25°C or more). While read-across is usually not possible, interpolation within homologous series of substances may be possible.

Knowledge on the melting point of a substance allows designating its physical state (liquid or solid), which is relevant for assigning physical hazard classes. Furthermore, it can be used to predict the partition behaviour of a substance between solid and gas phases, where a higher melting point indicates greater intermolecular forces and therefore less vapour pressure (Little Pro 2016). Solids with high-melting temperatures generally also have a poor solubility. The melting point thus serves as an indicator for the distribution of the substance within and between environmental media, such as water, soil and air, but also plays a role in the bioavailability of a substance in humans/mammals (National Research Council 2014, ECHA 2016, Little Pro 2016).

3.1.3 Boiling point

The boiling point is defined as the temperature at which the vapour pressure of a liquid equals 101.3 kPa, which means that the substance is completely gaseous at that temperature. Methods suitable to determine the boiling point experimentally are described in OECD TG 103 (OECD 1995). Under REACH, boiling point testing is not required for gases, or for solids that either melt above 300 °C or decompose before boiling. For borderline cases, boiling point tests should be conducted, but under reduced pressure (down to 0.2 kPa) if possible. Application of (Q)SARs

for the determination of the boiling point is discouraged, except when the mean absolute error of the method is lower than 2 K. While read-across is usually not possible, interpolation within homologous series of substances may be possible.

Knowledge on the boiling point of a substance allows designating its physical state (gas or liquid), which is relevant for assigning physical hazard classes. The boiling point has correlations with vapour pressure (paragraph 3.1.4) where higher boiling points indicate lower volatility. As such the boiling point indicates whether a substance will form flammable/explosive vapour-air mixtures, which is needed for the classification of flammable liquids, but also whether a substance will be available for inhalation as a vapour. The latter is particularly of relevance for human health risk assessment. The boiling point thus serves as an indicator for the distribution of the substance within and between environmental media, such as water, soil and air, but also plays a role in the bioavailability of a substance in humans/mammals (National Research Council 2014, ECHA 2016, Little Pro 2016).

3.1.4 Relative density

Relative density is defined as the quotient of the mass m and the volume V of a substance. Under REACH, relative density testing does not need to be conducted for substances that are gases at room temperature (estimation based on molecular weight and the Ideal Gas Laws is sufficient) and substances that are only stable in solution in a particular solvent and the solution density is similar to that of the solvent (indication whether the solution density is higher or lower than the solvent density is sufficient). Relative density should be determined experimentally, as is practically always technically feasible, (Q)SAR approaches are generally not applicable, and read-across is usually not possible. Interpolation within homologous series of substances may be possible though.

Knowledge on the relative density of a substance helps to understand the environmental distribution of the substance within and between water, soil and air. For gaseous materials, relative density is of value in determining the tendency to settle or to disperse when discharged at high concentrations into the atmosphere. For insoluble liquids and solids, (absolute) density will be a determining factor in the settling of the substance.

3.1.5 Vapour pressure

The vapour pressure of a substance is defined as the saturation pressure above a solid or a liquid substance at constant temperature. At the thermodynamic equilibrium, the vapour pressure of a pure substance is a function of temperature only. As such, the vapour pressure of a substance is the measure of a compound’s volatility. Methods suitable to determine the vapour pressure experimentally are described in OECD TG 104 (OECD 2006). Under REACH, vapour pressure testing does not need to be conducted if the melting point of a substance is above 300 °C, or if the standard boiling point is below 30 ºC. If the melting point is between 200 °C and 300 °C, a limit value based on measurement or a recognized calculation method is sufficient. Vapour pressure should be determined experimentally, but if that is not possible (Q)SAR approaches may be

used. While read-across is usually not possible, interpolation within homologous series of substances may be possible.

Knowledge on the vapour pressure of a substance allows its characterisation, which is relevant for assigning physical hazard classes. The vapour pressure is correlated to the melting and boiling points of a substance, and as such is considered a key parameter for evaluating the hazard and risk potential of substances to human health and the environment. In fact, vapour pressure is a pre-requisite for animal and environmental studies. It indicates whether a substance may be available for inhalation as a vapour, whether it may form flammable/explosive vapour-air mixtures, whether occlusive conditions are required for dermal studies, but also whether precautionary measures, e.g. respiratory protection equipment, are needed to minimize emission to air and occupational exposure. As such it plays an important role in human health risk assessment. The vapour pressure also allows determination of the relative volatility of a substance from aqueous media and soil, in terms of the Henry's Law constant (HCL) and partition coefficient air/soil.

The HCL is a key parameter in determining the environmental fate of substances, where it is used to quantify the partitioning of substances between the aqueous phase and the gas phase such as rivers, lakes and seas with respect to the atmosphere. It is also used in human health risk assessment as indicator for human/mammalian alveolar absorption.

The HCL can be experimentally determined using a dynamic or static equilibration approach, calculated as the ratio of water solubility (C_w) to vapour pressure (vp), or estimated by QSAR methods. The HCL indicates how readily a substance evaporates. If the HCL is high (~100 $Pa \cdot m^3/mol$) evaporation occurs rapidly. For such substances volatilization to air is an extremely important removal process from water, but it also results in distribution of the substance over a large area. Substances with low HCL values (<0.01 $Pa \cdot m^3/mol$) tend to persist in water as they are less volatile than water. Such substances also often adsorb onto organic carbon in soil or sediment (ECHA 2016). In general substances with a HCL below 0.1 $Pa.m^3/mol$ are considered difficult to test (OECD 2002). I should be noted though that care should be taken when using the HCL for evaluation of the removal of a substance from the aqueous phase, as the HCL value of a substance depends on the environmental parameters for the specific water bodies in question, such as water characteristics (e.g. pH, dissolved salts, dissolved organic matter, etc.), the depth and the gas exchange coefficient (influenced e.g. by wind speed and water flow rate).

3.1.6 Surface tension

The surface tension of a substance is defined as the free surface enthalpy per unit of surface area (European Commission 2008). Methods suitable to determine the surface tension experimentally are described in OECD TG 115 (OECD 1995). Surfactants are surface active substances that are capable of reducing the surface tension of water, of forming monolayers at the water-air interface, and of forming emulsions and/or microemulsions and/or micelles in water. The maximum concentration of the freely solubilised surfactant in water is referred to

as critical micelle concentration (OECD 2002). Surfactants are defined by a surface tension that is below 60 mN/m (ECHA 2014). Under REACH, surface tension only needs to be determined if based on structure, surface activity is expected or can be predicted, or surface activity is a desired property of the material. If the water solubility of a substance is below 1 mg/L at 20 °C the test does not need to be conducted. Currently there are no reliable (Q)SAR methods available for sufficiently accurate predictions of surface tension. While read-across is usually not possible, interpolation within homologous series of substances may be possible.

Knowledge on surface tension is important since decreasing the surface tension of water may impact on the properties of the solution and other physicochemical measurements. It also important as performing fate, e.g. bioaccumulation, and effect studies with surfactants is difficult due to their intrinsic tendency to form an aqueous dispersion or emulsion. Furthermore, most surfactants exhibit at least to some extent toxicity to aquatic organisms as they react with the biological membranes of the organisms.

3.1.7 Water solubility

The solubility of a substance in water is specified by the saturation mass concentration of the substance in water at a given temperature. An easier to interpret definition is given by OECD TG 23 with water solubility being the maximum attainable concentration (in pure water) or the concentration at thermodynamic equilibrium between aqueous phase and solid (or liquid or gaseous) pure phase (OECD 2002). Water solubility can vary considerably with temperature and pH, and it is therefore important that the solubility data are reported with the experimental conditions. Water solubility testing should almost always be possible, and should be determined experimentally, e.g. shake flask or column elution method, as described in OECD TG 105 (OECD 1995). Under REACH, a water solubility test does not need to be conducted if the substance is hydrolytically unstable at pH 4, 7 and 9 (half-life < 12 hours), or if the substance is readily oxidisable in water. If the substance appears 'insoluble' in water, a limit test up to the detection limit of the analytical method is sufficient. Good estimations can be made with QSAR models (e.g. WSKOW, WATERNT (US EPA 2012). While read-across is usually not possible, interpolation within homologous series of substances may be possible.

Knowledge of the water solubility of a substance is of paramount importance for several reasons. The water solubility of a substance determines to a great extent its mobility, with water soluble substances being more likely distributed by the hydrological cycle and taken up by humans and other living organisms. As such it is of importance for modelling distribution and exposure, but also a prerequisite for setting up test conditions for fate and effect studies. The latter is especially important for substances that are poorly water soluble (<100 mg/L) and for which it is challenging to achieve, maintain and measure exposure concentrations (OECD 2002), and for substances that have a very low water solubility (i.e. in the low µg/L range) and for which aquatic toxicity testing might not even be feasible, and terrestrial toxicity testing should be preferred. As such water solubility, can serve as a regulatory trigger.

Furthermore, water solubility is a key input parameter for various QSAR models, and is required to derive environmental parameters such as K_{ow}, K_{oc} and Henry's Law Constant.

3.1.8 Dissociation constant

Dissociation is the reversible splitting of a substance into two or more chemical species, which may be ionic (OECD 1981). For substances that have only one dissociating group, the acid dissociation constant (pK_a) is equivalent to the pH at which the ionised and non-ionised forms are present in equal concentration. For substances with more than one dissociating group, pK_a values can be determined for each dissociation step. Method to determine the pK_a value(s) of a substance is described in OECD TG 112 (OECD 1981). Under REACH, the study does not need to be conducted if the substance is hydrolytically unstable (half-life less than 12 hours) or is readily oxidisable in water, or it is scientifically not possible to perform the test for instance if the analytical method is not sensitive enough.

Knowledge on the dissociation behaviour of a substance is important, as the pK_a governs the ionization state of a substance. The physicochemical properties, e.g. water solubility and partition coefficients (K_{ow}, K_{oc}), of dissociated and non-dissociated forms can be markedly different, which in turn affects the substance behaviour, bioavailability and transport. This is especially evident when the pK_a value is within the environmentally-relevant pH range of pH 4 to 9. In human health risk assessment this means that the ADME (absorption, distribution, metabolism, and excretion) properties of a substance can be altered, while for environmental risk assessment the bioaccumulation potential and adsorption behaviour can be strongly affected. Therefore, the pK_a of a substance is considered basic knowledge needed to correctly setup test conditions for fate and effect studies for human health and environmental risk assessment.

3.1.9 pH value

pH stands for 'potential of Hydrogen' and is a measure of acidity (pH <7) or alkalinity (pH >7) of water-soluble substances that is calculated as the negative logarithmic value of the Hydrogen ion (H^+) concentration.

pH is basic physicochemical knowledge that should be considered before any toxicology studies are to be conducted. Depending on the acidity or alkalinity of the substance, the pH is determined by titration with a standardized strong base or acid, respectively, as described in OECD TG 122 (OECD 2013). Strong acids (pH≤2) or bases (pH ≥11.5) are corrosive to bio-membranes and do damages to tissues. Such corrosive substances are naturally acutely toxic and should, in accordance with current EU and OECD guidelines, not be tested on animals for skin corrosion, eye irritation, sensitisation and 28d repeated dose toxicity studies. In the absence of other information, substances with a pH of ≤2 or ≥11.5 should be considered as Skin Corrosive Category 1 under GHS and, thus also under CLP.

3.1.10 *Partition coefficient n-octanol/water*

The *n*-octanol/water partition coefficient (K_{ow}) is defined as the ratio of the equilibrium concentrations of a dissolved substance in a two-phase system consisting of the largely immiscible solvents *n*-octanol and water. Typically the log K_{ow} is defined as the partition coefficient of the neutral, undissociated form of a substance, usually measured at 25°C and preferably within an environmentally relevant pH range of 5 to 9 (ECHA 2016). For pharmaceuticals often a D_{ow} is determined at pH 7.4, which is of physiological relevance, regardless of the dissociation state of the molecule. The degree of dissociation of a substance, and thus the pH at which is tested, can have a marked effect on the log K_{ow}. Under REACH, log K_{ow} cannot be waived, with the exception of inorganic substances. For all other substances, including difficult to test substances such as poorly soluble, volatile, surface active, ionisable, or rapidly decomposing substances, or mixtures of substances, a log K_{ow} value is required. If experimental testing using the shake flask (OECD TG 107 (OECD 1995)), slow stirring (OECD TG 123 (OECD 2006)), or HPLC (OECD TG 117 (OECD 2004)) methodology is not possible, and a log K_{ow} cannot be estimated from the individual solubilities (as is done for surfactants), a log K_{ow} value can be calculated based on the molecule's structure.

Knowledge on log K_{ow} is of paramount importance for both human health and environmental risk assessment, for several reasons. It allows predicting the distribution of a substance between environmental compartments. Generally spoken, substances with high log K_{ow} values tend to adsorb strong to organic matter in sediment and soils (having high K_{oc} values) and tend to bioconcentrate in biota due to a strong affinity for lipids. Exceptions are metals, surface active compounds, and organic substances that have a strong affinity for proteins, e.g. perfluorinated compounds. For the vast majority of organic compounds though, log K_{ow} >4.5 is a valid descriptor for bioaccumulation potential, and a justified trigger for bioconcentration testing in aquatic organisms. Log K_{ow} is also an important determinant of human/mammalian oral and skin bioavailability, which is directly related to higher exposure levels and toxicity. Finally, log K_{ow} is an indicator for toxicity through a narcosis mode of action for non-ionic organic substances that is associated with altered structure and function of cell membranes. Considering above, log K_{ow} is an important input parameter in numerous estimation models and algorithms for environmental partitioning, sorption, bioavailability, bioconcentration, bioaccumulation and also human toxicity and ecotoxicity.

3.1.11 *Granulometry*

Granulometry is not a specific physicochemical property of a substance, rather particle size distribution is defined by the manufacturing process, and subsequent uses. The following terms are relevant with respect to granulometry: Particle, which is a minute piece of matter with defined physical boundaries; Fibre, which is a water insoluble particle with an aspect ratio (length/diameter > 3) and diameter < 100 µm; Agglomerate, which is a collection of weakly bound particles of aggregates or mixtures of the two where the resulting external surface area is similar to the sum of the surface areas of the individual components; and Aggregate, which is a particle comprising strongly

bonded or fused particles where the resulting external surface area may be significantly smaller than the sum of calculated surface areas of the individual components. Methods suitable to determine the particle size distribution, fibre length and diameter distributions are described in OECD TG 110 (OECD 1981). Under REACH, granulometry does not need to be conducted if a substance is marketed or used in a non-solid or granular form. When describing particle size distribution, measures can represent the actual diameter of a particle, e.g. an optical, measured or geometric diameter, or can reflect a relative diameter, e.g. aerodynamic or particle diffusion diameter. The latter are used to compare particles of different sizes, shapes and densities. There are no QSARs available for granulometry and grouping and read-across approaches are not considered applicable.

It should be noted that nanomaterials, which are very small particles, can differ in their behaviour and toxicity compared to the same substance tested in solid form. Therefore, additional parameters have been specified for nanomaterials in an amendment to REACH Guidance documents R7a and 7c (draft version; (ECHA 2016)), i.e. surface chemistry, size, shape, surface area, solubility (in relevant biological fluids/testing media), dispersability (refers to the relative number or mass of particles in a suspending medium, and relates to stability (Sellers, Deleebeeck et al. 2015), aggregation and agglomeration in relevant media, and is dependent on e.g. van der Waals energy, Hamaker constant, zeta potential), dustiness, biological reactivity (reactive oxygen species (ROS) production), photoreactivity, stability in storage and rigidity for fibres.

Knowledge on particle size distribution is relevant as solids in fine powder or dusty form readily enter the air, which can lead to combustible/explosive dusts – air mixtures, but also changes the route of exposure of humans. In human health risk assessment, the aerodynamic diameter is used to predict where in the respiratory tract particles may be deposited (e.g. inhalable, thoracic and respirable fractions correspond to aerodynamic diameters of >100, 10 and 4 µm, respectively). This is relevant, as human toxicity is related to the place of deposition into the respiratory tract. Furthermore, the dustiness, i.e. capacity to produce airborne dust, of materials is often considered, especially when assessing quality of the workplace atmosphere. Particle size distribution also plays an important role in environmental risk assessment, as it affects the distribution of the substance in the environment, but also determines the uptake efficiency by organisms.

3.1.12 Adsorption/desorption

Adsorption is not a specific physicochemical property of a substance, rather it indicates the binding capacity of substances to solid surfaces, which can be reversible or permanent depending on the type bond (i.e. due to van der Waals interactions, hydrogen bonding to hydroxyl groups, ionic interactions, covalent bonding, etc.). Soil sorption is described by the organic carbon normalized adsorption coefficient (K_{OC}), which is the ratio of a substance concentration sorbed in the organic matter component of soil or sediment to that in the aqueous phase at equilibrium. Under REACH information on adsorption (and desorption) is required, except if based on the physicochemical properties it can be

expected that a substance will have a low potential for adsorption, or that the substance and its relevant degradation products will decompose rapidly. Low potential for adsorption is generally considered a log K_{ow} below 3, but this cut-off value is not always a good predictor, i.e. substances that are water soluble can have a low log K_{ow} and still adsorb, the log D_{ow} and adsorption potential of ionising substances will be pH-dependant and for surface active substances the log K_{ow} is difficult to determine at best (see section 3.1.10). Under REACH, for low tonnage substances (>10 t/y) an initial screening based on read-across or QSAR estimates of K_{OC} is recommended. If a risk is identified for the sediment/soil compartments, or if it concerns a higher tonnage substances (>100 t/y) additional information on sorption is needed. The K_{OC} can estimated by the HPLC method (OECD TG 121 (OECD 2001)), but for definitive data the batch equilibrium method (OECD TG 106 (OECD 2001)) is preferred, unless the substance is unstable in which case the column leaching method (OECD TG 312 (OECD 2004)) can be used. For human medicinal products, QSAR or HPLC estimated K_{OC} values are not considered acceptable for screening, and determination using the batch equilibrium method is required for both soil/sediments (OECD TG 106) and sewage sludge (OPPTS 835.1220 (US EPA 1996)).

Knowledge on the adsorption and desorption behaviour of a substance is of great importance for environmental risk assessment for several reasons. Firstly, it determines to a great extent the environmental distribution of a substance, i.e. adsorption to sewage sludge and suspended solids will lower concentrations in surface water while increasing them in soil and sediment, while desorption form soil will mobilise substances after which they leaching into groundwater. These routes can be modelled, as is discussed in more detail in section 4.2. Secondly, it determines the availability of a substance, i.e. if a substance is highly sorptive toxicity testing in aquatic systems might not be feasible, and testing in soil and/or sediment systems will be required. If the substance binds irreversible it might not be bioavailable to soil and sediment inhabiting species, thus reducing the toxicity, but also to microorganisms, which will slow down degradation rates, and increase the persistence of the substance, which is relevant for PBT assessment.

3.1.13 Viscosity

Viscosity is the (inner) resistance of a substance (gas, liquid) to a shift caused by laminar flow. Dynamic viscosity expresses a fluids resistance to shearing flows, where adjacent layers move parallel to each other with different speeds. Kinematic viscosity is the ratio of the dynamic viscosity to the density of the fluid, and is as such a measure of the resistive flow of a fluid under the influence of gravity. It should be noted that this physicochemical property is particularly temperature dependent, with increased temperatures leading to reduced viscosity. Under REACH, viscosity needs to be determined, unless the substance is a solid, or the liquid is an explosive, pyrophoric or self-reactive substance. Different standardized methods are described in OECD TG 114 (OECD 2012). Experimental values determined at least two different temperatures are preferred. QSAR estimations can be used, but are generally discouraged for the purpose of classification/risk assessment. Read-across is not possible.

Knowledge on viscosity is needed for substance characterization and classification of aspiration hazards of liquids (kinematic viscosity < 20.5 mm^2 at 40°C). Generally, liquids with lower viscosity are more hazardous as they can spread more quickly. Furthermore, knowledge on a substance viscosity indicates how readily a substance will penetrate within soil.

3.1.14 Electronic properties

Electronic properties as such have not been specified under REACH as physicochemical parameters that need to be reported. Nevertheless, they are considered relevant with regard to toxicity, and are included in this overview. The frontier orbital energies HOMO (=highest occupied molecular orbital) and LUMO (=lowest unoccupied molecular orbital) reflect the chemical reactivity (ΔE energy gap between HOMO and LUMO) of a substance with nucleophiles and electrophiles, which translates to reactivity with biomolecules *in vivo*. This is directly related to toxicity in aquatic species and humans/mammals (see Figure 3). Furthermore, molecular electronic dipole moments (μ) and dipole polarizabilities (α), are important in determining the energy, geometry, and intermolecular forces of molecules, and are often related to biological activity. These properties are often estimated with quantum mechanic calculations based on the structure of the substance.

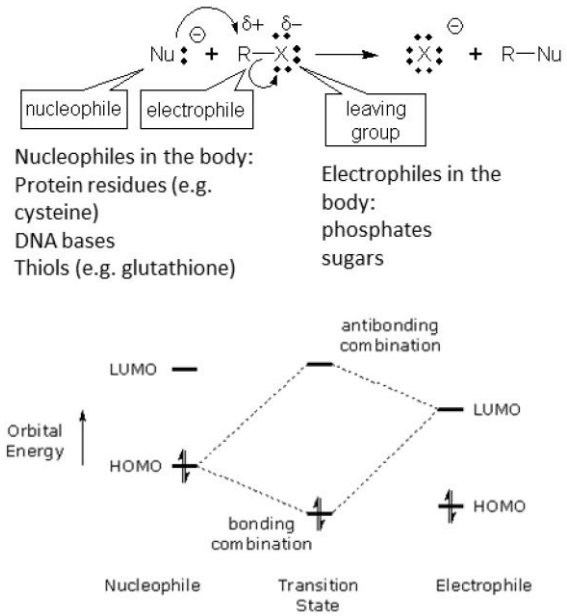


Figure 2 Relation of frontier molecular orbital energies to covalent interactions of nucleophiles and electrophiles, as illustrated with a generic nucleophilic substitution reaction (obtained from National Research Council (2014)).

3.1.15 Other physicochemical properties important for physical hazard assessment

In addition to the above discussed physicochemical properties, several other properties are important for substance characterization, classification of substances and assigning of physical hazard classes according to CLP regulation:

- Flashpoint: the lowest temperature at which a chemical can vaporize to form an ignitable mixture in air. A lower flash point indicates higher flammability.
- Flammability: ability of a substance to burn or ignite, causing fire or combustion.
- Explosiveness: ability of a solid or liquid substance to cause by itself a chemical reaction of producing gas at such a temperature and pressure and at such a speed as to cause damage to the surroundings. Pyrotechnic substances are included even when they do not evolve gases.
- Self-ignition temperature: the lowest temperature at which a substance will ignite when mixed with air under defined test conditions.
- Oxidising properties: ability of substance, while in itself not necessarily combustible, to cause or contribute to the combustion of other material by generally yielding oxygen. Oxidizing gases are further specified as being oxidizing than air.

3.2 Use of physicochemical properties in human health risk assessment

As discussed in the sections above, physicochemical properties can be directly linked to physical hazards classes specified under the GHS and CLP regulations. From a human health perspective, the most relevant undesired physical hazards are flammability, gases under pressure, explosiveness, oxidizing liquids and corrosivity to metals. Generally, for substances classified as such, the highest risk is observed during manufacturing and transport phases, while at later phases, i.e. post-manufacturing, consumer use and end-of-life, the hazard is often less pronounced or, in some cases, absent altogether (National Research Council 2014). In addition to physical hazard classes, one of the health hazard classes, i.e. Skin Corrosive Category 1, is also directly linked a physicochemical property, i.e. the pH value, with strong acids (pH≤2) and bases (pH ≥11.5) being corrosive to bio-membranes and thus leading to tissue damage.

Next to using knowledge on physicochemical properties to classify substances, such knowledge is also a prerequisite to correctly design and perform toxicity studies with humans and/or other mammals, but also to predict substance toxicity. The influence of individual physicochemical property will not be repeated here (see sections above), instead this section aims to provide a brief and general overview of how physicochemical properties are used in human health risk assessment. More details can be found in subsequent chapters. Basically, all aspects of human health risk assessment are influenced by the physicochemical properties of a substance. Distribution, human exposure and toxicokinetic behaviour can be estimated using physicochemical properties as input in models (see respective chapters for details). Estimation of toxicodynamic properties, i.e. the interaction between a substance and it biological targets and subsequent (toxicological) effects, on the other hand is hampered by the lack of apparent direct links to specific physicochemical properties.

Substance's molecular size and shape, lipophilicity and hydrophobicity (expressed by log K_{ow}), ionization potential (expressed by pK_a) and hydrogen bonding capacity (which in itself influences a great deal of physicochemical properties, including melting/boiling points, vapour pressure, water solubility, ionization, and adsorption (Ferguson 1956)) are major determinants of toxicokinetic behaviour (National Research Council 2014). Modelling using these parameters as input allows one to get a better understanding of important processes in humans and/or mammals, such as the rate of absorption of a substance into the bloodstream, its distribution to the organs and tissues, and its rate of elimination (clearance) by means of metabolism and/or excretion (see chapter 5). These physicochemical properties, together with vapour pressure, also determine the bioavailability of substances and the route by which mammals will be exposed, i.e. by means of ocular, oral, respiratory and dermal exposure. Property limits have been set that are associated with an increased bioavailability through one of these four routes (Table 1).

Table 1 Combinations of property limits associated with increased bioavailability through the four main routes of exposure in mammals (obtained from National Research Council (2014))

Exposure route	**Physicochemical Property**	**Property Limit**
Ocular	Water solubility	Variable
	Molecular size	< 500 Da (corneal epithelium) < 10000 Da (conjunctival epithelium)
	Vapor pressure	< 0.0001 mm Hg
Oral	Molecular size	< 500 Da
	Log K_{ow}	0-5
	Non-ionized at Gastrointestinal tract pH	-----
Respiratory (Lungs)	Particle size	< 5 µm
	Molecular size	< 400 Da
	Vapor pressure	< 0.0001 mm Hg
Dermal	Molecular size	< 400 Da
	Log K_{ow}	0-6
	Presence of solvents	-----
	Ionization (polar, ionized)	-----

3.3 Use of physicochemical properties in environmental risk assessment

The aim of this section is to provide a brief and general overview of how physicochemical properties are used in environmental risk assessment. Details are provided in the respective chapters below. Physical hazard classification plays a far less pronounced role in environmental risk assessment, with the aim primarily being protection of consumers and workers. In environmental risk assessment, knowledge on physicochemical properties is used to correctly design and perform environmental fate and toxicity studies, and to predict distribution,

exposure scenarios, bioaccumulation potential and toxicity of a substance.

The molecular weight, water solubility, the ionization potential (expressed by pK_a), the lipophilicity (expressed by log K_{ow}), the adsorption potential (expressed by K_{OC}) and the physical state of a substance (governed by the melting point, boiling point and vapor pressure) determine to a great extent the partitioning of a substance over the different environmental compartments, e.g. air, water and soil, and the biota living in these compartments. These are, except for the K_{OC}, the same parameters that are also relevant for predicting exposure of humans/mammals and toxicokinetics. Details are given in the sections above for the individual parameters, but it is clear that substances with a high HLC (depending on vapour pressure and water solubility) will, due to their volatility, escape from soil or water and primarily be present in the air. Conversely, substances with low water solubilities, and a high propensity to sorb onto organic carbon (high K_{oc}) or move into lipid phases (high log K_{ow}) will likely remain in soils or sediments or move into biota, respectively. In aquatic species, bioavailability is positively correlated with log K_{ow}, and influenced by aqueous solubility, molecular size, and ionization. Generally, it can be said that substances that are not bioavailable are likely to have low toxicity, while high bioavailability is not an indication of high toxicity. In the table below physicochemical property limits are listed that are known to favour reduced acute and/or chronic aquatic toxicity.

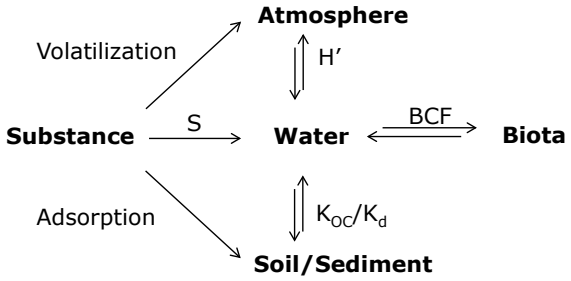


Figure 3 Partitioning of a substance over different environmental compartments and biota.

The bioconcentration factor (BCF) is either experimentally determined (see section 5.2), or QSAR estimated using the log K_{ow}.

Table 2 Changes in physicochemical properties to favor reduced aquatic toxicity (adapted from National Research Council (2014))

Physicochemical Property	**Changes**
Molecular size and weight	Generally, as molecular weight increases, aquatic bioavailability and toxicity decrease. At MW > 1000 Da, bioavailability is negligible. Caution must be taken, however, to consider possible breakdown products that may have MW < 1000 Da and exert toxicity.
Log K_{ow}	Log K_{ow} usually correlates exponentially with acute aquatic toxicity by narcosis for non-ionic organic chemicals up to a value of about 5-7. Chemicals with K_{ow} < 2 have a higher probability of having low acute and chronic aquatic toxicity (Voutchkova, Osimitz et al. 2010, Voutchkova, Kostal et al. 2011). For ionisable organic chemicals, log $D_{ow, 7.4}$ < 1.7 has been suggested as having an increased probability of being safe to freshwater fish than those with log $D_{ow, 7.4}$ > 1.7 (Kostal, Voutchkova-Kostal et al. 2015).
Water solubility	Generally, compounds with higher Log K_{ow} have lower water solubility. Very poorly water-soluble chemicals (<1 µg/L) generally have low bioavailability and are less toxic
ΔE energy [HOMO-LUMO]	The ΔE reflects broad chemical reactivity. It was recently reported that chemicals with ΔE > 9 eV (as calculated by semi-empirical methods) or > 6.5 eV (as calculated by DFT) are much less likely to be acutely or chronically toxic to aquatic species (Voutchkova, Kostal et al. 2012, Kostal, Voutchkova-Kostal et al. 2015).

3.4 Opportunities for integration of human health and environmental risk assessment with regard to physicochemical and solvation properties

Physicochemical and solvation properties are intrinsic properties of a substance, and as such are not dependent on the protection goal of a risk assessment framework, i.e. human health or the environment. Most properties are important for both risk assessment frameworks, including the molecular size and shape, the lipophilicity and hydrophobicity (expressed by log K_{ow}), the water solubility (expressed by C_w), the acidity or alkalinity (expressed by pH), the ionization potential (expressed by pK_a), and the vapour pressure (vp) of a substance. There are of course also physicochemical properties whose importance is more pronounced in one of the risk assessment frameworks. Examples are the adsorption potential (expressed by log K_{oc}) that is primarily used in the ERA to determine environmental distribution of a substance, and properties such as flashpoint, flammability, explosiveness, self-ignition temperature and oxidising properties that are used to assign physical hazard classes and as such are more important for HRA. Overall, experimental determination of physicochemical and solvation properties

is rather standardized, and the QSARs for estimating physicochemical and solvation properties can be used in both RA frameworks, demonstrating that integration has already been achieved with regard to generating data for physicochemical and solvation properties.

The usage of physicochemical properties to correctly design and perform studies, as well as, input for *in silico* approaches, e.g. QSARs, distribution, exposure and/or metabolism models, is tailored to the needs of the respective RA framework. That said cross-fertilization takes place and efforts are made to translate *in silico* approaches and methods from HRA to ERA and vice versa with respect to exposure assessment (Chapter 4), toxicokinetics & dynamics (Chapter 5), and toxicology (Chapter 0). These approaches are discussed in more detail in the respective chapters.

Overall, it can be concluded that with regard to physicochemical and solvation data integration has already been achieved, and that further integration is not expected to substantially improve risk assessment or increase the overall efficiency.

4 Exposure assessment

Exposure assessment is the process of considering and estimating the extent, frequency and duration of (intentional or unintentional) emissions of substances into the environment, resulting in exposure of human and ecological receptors, which together encompass all levels of biological organisation, i.e. cells, tissues, organisms, populations, communities, and ecosystems. Exposure assessment can be considered to consist of two parts, the first part focussing on the emission and distribution of substances over different environmental compartments and biota (see section 4.1), while the second one determines the exposure of man and other organisms to substances (see section 4.2). It should be noted that exposure can also occur to biological and physical agents, e.g. noise, radiation, microbial hazards, as well as social stressors, but these are not considered here, because of their unique characteristics. This chapter will give an overview of the current practice in human and environmental exposure assessments of substances, which are in essence linked, since both need to account for all interacting compartments as well as the physicochemical properties of a substance. Nevertheless, as different protection goals are pursued, human and environmental exposure assessment developed and used in parallel their own data, methods, scenarios and models. The efforts that have already been made to integrate human and environmental exposure assessment will be discussed, with the aim to identify if further integration is wanted and/or is feasible.

4.1 Emission and distribution

Release of substances into the environment, eventually leading to exposure of humans and other organisms via the environment, can occur during production, formulation, use and disposal phases from both point and diffuse sources. Emissions can vary in relation to level, time and place (e.g. continuous versus block or peak emissions), and will greatly depend on the type of substance and its intended usage, examples being: a plant protection product or biocide that is intentionally sprayed in an open field; human and veterinary pharmaceuticals that upon excretion enter the surface water after respectively wastewater treatment or directly; industrial chemicals that are used as closed intermediates and that under normal conditions are not emitted; substances added to food or feed or cosmetics that are applied directly to humans which are widely dispersed. These are just a few examples of possible uses, but it is clear that some routs can be specific to human processes. Emissions from certain uses can be reduced or prevented by applying risk reduction measures to optimise substance flow and processes, e.g. recycling of waste and substitution of undesired substances, as well as end-of-pipe measures, e.g. incineration and wastewater treatment.

Once a substance has reached the environment its physicochemical properties, i.e. vapour pressure, water solubility, molecular weight, octanol-water partition coefficient, melting point and biodegradability, will determine to a great extent its fate and distribution within an

environmental compartment, e.g. soil, sediment, water and air, or between several compartments and biota. The concentrations of the substance can be measured in the different compartments, but they represent single measurements in time and space. Unless monitoring data is obtained at regular time intervals and from different locations, their use is limited for risk assessment, with the exception of measurements in pristine environments that indicate that a substance is persistent and that long-range transport occurs. In risk assessment, environmental release and distribution is therefore often modelled using the physicochemical properties of a substance either for the separate compartments or in multimedia models, e.g. Mackay level III model, where the different processes are considered simultaneously (see Figure 5), on different spatial scales, i.e. on a local scale in the vicinity of point sources, such as industrial sites and wastewater treatment plants, or on a regional scale which includes a larger area consisting of all point sources in that area as well as dispersive sources. Under REACH EUSES, into which Simple Treat and Simple box have been incorporated, is used to model fate and distribution of a substance, while in the plant protection products framework FOCUS, PEARL and PRZM are used to determine distribution between compartments. It should be noted that point of entry can be decisive in the final distribution of a substance, and thus its chemical fate.

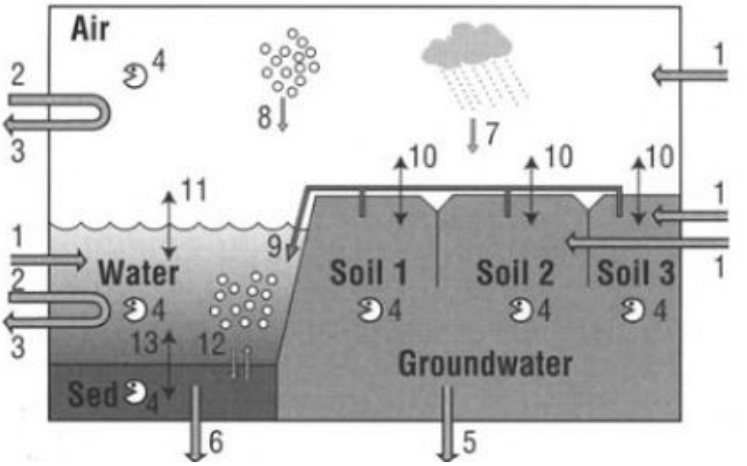


Figure 4 Diagram of a multimedia mass balance model concept (obtained from van de Meent and de Bruijn (2007)).

Following processes are specified: (1) Emission; (2) Import; (3) Export; (4) Degradation; (5) Leaching; (6) Burial; (7)Wet deposition; (8) Dry aerosol deposition; (9) Run-off; (10 & 11) Gas absorption and votalization; (12) Sedimentation and resuspension; (13) Sorption and desorption.

4.1.1 Degradation route

One of the processes that profoundly impacts a substances' distribution is its degradation. Degradation can be abiotic, e.g. hydrolysis in water and photolysis in the atmosphere, but also biotic, e.g. microbial biodegradation and metabolism in higher organisms. Microbial biodegradability is often the driving disappearance process in the environment, and can be estimated using physicochemical data only

(e.g. BIOWIN), but it is preferably generated experimentally. Initially the ready biodegradability of a substance is determined, where the degradability of a substance in diluted sewage sludge is assessed. If the substance does not degrade sufficiently within 28 days (60 or 70% depending on the OECD test guideline applied), the substance is considered not ready biodegradable, and degradability of the substance under more favourable conditions can be assessed where higher amount of sludge and lower amount of test substance is applied. These studies only serve to see if the substance will degrade, but cannot be used to determine the speed at which degradation occurs. Simulation studies in surface and/or marine water (OECD TG 309), sediment (OECD TG 308) and soil (OECD TG 307) can be conducted to determine the speed at which degradation occurs. Assuming first-order degradation kinetics a degradation half-life (DegT50) for the different compartments can be calculated. It should be noted that there are differences between frameworks, as for plant protection products often a dissipation half-life (DissT50) is reported that takes degradation, but also sorption and evaporation into account, while under REACH a distinction must be made between degradation and other processes. Another difference between frameworks is the temperature that is used to determine the speed of degradation. Using the Arrhenius equation DT50 values can be recalculated to other temperatures, but especially when the persistence of a substance is considered, these differences in interpretation can play a significant role.

4.1.2 Adsorption route

Substances can also become unavailable to biota by other processes than degradation, i.e. by adsorption to organic matter or fine clay particles in soil and/or sediment. This is expressed by the adsorption coefficient (log K_{oc}). The binding can be reversible, but substances can also sorb very strongly forming non-extractable residues (NERs), which makes them practically unavailable to biota under the prevailing environmental conditions. However, upon ingestion these substances can become available due to lower pH values. If the substances are covalently bound to soil particles or are incorporated into biomass (biogenic fixation) the substances ordinarily do not become available. The distinction between these different types of NERs is rather complex, and the interpretation of NER in degradation assessment is the topic of current research (e.g. (Kästner, Trapp et al. 2018, ECHA 2019)).

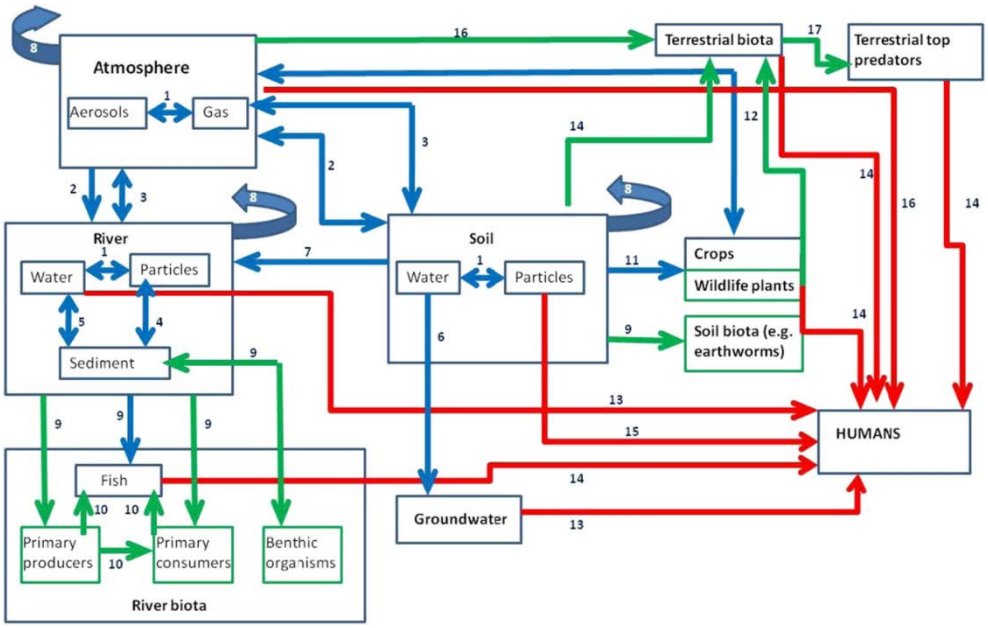
4.1.3 Bioconcentration route

Distribution of substances is also affected by the potential of a substance to bioaccumulate in organisms, leading to increased internal exposure concentrations. This route is important for both human and environmental risk assessment, and is discussed in detail in Chapter 5 that deals with toxicokinetics. In short, it can be said that lipophilic substances (expressed by high log K_{ow}) tend to bioaccumulate in lipid tissue, unless their (large) molecular size prevents crossing over the cell membrane. There are also substances, e.g. poly- and perfluorinated substances, where bioaccumulation is mediated by protein binding instead of lipid accumulation. In most frameworks, bioconcentration in aquatic organisms is assessed as part of the ERA. This is preferably conducted in fish (OECD TG 305), but bioaccumulation can also be assessed in aquatic (OECD TG 315) or terrestrial invertebrates (OECD

TG 317). Values derived from invertebrates tend to be more conservative, as metabolism is limited compared to vertebrate species. In HRA, bioaccumulation is considered in mammals, by measuring the substance over time in various tissues and blood. Field measurements can also be informative with regard to bioaccumulation potential in humans, i.e. half-lives of substances in human blood can be derived from human populations that have been exposed to substances over a prolonged period of time.

4.2 Exposure

In environmental risk assessment, exposure via the different environmental compartments shown in Figure 5 is taken into account by calculating the predicted environmental concentrations (PECs) for the relevant compartments. The PECs can be calculated as a local or regional PEC, and after single ($PEC_{initial}$) or after multiple ($PEC_{plateau}$) applications. There are differences between frameworks in PEC calculations to account for the differences that exist in usage and exposure scenarios. But in all frameworks the PEC for the surface water compartment is calculated first, and is subsequently used to calculate the PEC for the groundwater compartment. If based on physicochemical properties of the substance a risk to the sediment or soil compartment is anticipated (high log K_{OW} and/or high K_{OC} and/or significant shifting to sediment), than a PEC for soil and/or sediment compartments is calculated using the equilibrium partitioning method. PEC calculation in lower tiers is based on worst-case assumptions, e.g. a total residue approach is followed assuming no metabolites/ transformation products are formed. In higher tiers refinements can be conducted based on metabolism or degradation, but if substances are formed in significant amounts (10%) they should be assessed too. Besides direct exposure, organisms in higher trophic levels can be indirectly exposed to substances that bioaccumulate in their prey. Under the biocide, REACH, veterinary medicinal products and plant protection products regulations secondary poisoning is assessed to account for the exposure indirect route (see section 5.2).



Blue lines represent common pathways, green and red lines pathways specific to environmental and human exposure assessment, respectively; Following processes are specified: (1) partition; (2) dry and wet deposition; (3) absorption/volatilization; (4) deposition/resuspension; (5) diffusion; (6) leaching; (7) runoff and erosion; (8) degradation; (9) direct uptake; (10) food-web transfer; (11) root uptake; (12) interception by leaves and diffusive exchanges; (13) drinking water; (14) food ingestion; (15) soil particles ingestion; (16) inhalation; (17) food web relationships.

Figure 5 Main environmental pathways to be considered in environmental and human exposure assessment (obtained from Ciffroy, Péry et al. (2016)).

In human health risk assessment, the same exposure routes apply as for the environmental risk assessment, with the addition that there are human specific exposure pathways, corresponding to purely anthropogenic life conditions like occupational exposure, intentional use of products by consumers (e.g. cosmetics), exposure to indoor air, treated drinking water or food packaging contaminants, and urban way of life. Recently, US EPA published the draft 'Guidelines for Human Exposure Assessment' (Tulve, Olsen et al. 2016), which focusses on assessing exposure of adults, children and other vulnerable groups within the human population, to substances in a nonoccupational environment. The EPA document can be used as a general reference document, as it describes the principles of exposure science and assessment, provides guidance on the various approaches that can be used to conduct an exposure assessment and provides references for more detailed information. The EPA document is not intended as a detailed instruction manual though, nor does it endorse certain models, or does it deal with emerging topics, such as high-throughput exposure assessment, the implications of *in vitro* risk assessments on the field of exposure assessment or novel exposure models such as ConsExpo. In this chapter, the exposure to workers, as modelled by ConsExpo, will also not be further discussed, as this is specifically related to exposure for humans that integration with environmental exposure assessment is considered relevant. Human exposure can be acute or chronic, and can occur via different routes, i.e. orally by ingestion of food or water, respiratory, dermally or through their eyes (ocular). These patterns and

routes of exposure also apply to other organisms, but in environmental exposure assessment the route of entry is only separately determined for birds and mammals. With regard to oral consumption by humans, maximal daily intake doses are set for compounds, such as industrial chemicals and pharmaceuticals. For plant protection products, Maximum Residue Limits (MRLs) are set for chronic ingestion as fruits and plants can contain residues of applied plant protection products. In human risk assessment, in contrast to environmental risk assessment, attention is given to population characteristics and sociodemographic factors that might increase exposure or predispose a life stage, vulnerable group or population to greater risk. These factors can include age, sex, genetic susceptibility, cultural characteristics, behaviours, occupation, socioeconomic status, race/ethnicity and geographic location. Furthermore, not only exposure of populations is taken into account, but also groups and individuals.

4.3 Opportunities for integration of human health and environmental risk assessment with regard to exposure assessment

The first part of exposure assessment, i.e. the emission and distribution of substances, is rather comparable for both risk assessment frameworks, as models to predict substance distribution are based on common properties of chemicals. The difference between human and environmental exposure assessment becomes more apparent in the part where environmental concentrations are used to predict internal exposure concentrations. The latter predictions are largely based on inter-species extrapolations, and while being methodologically similar they do account for the specific routes of the respective framework. In recent years, considerable effort has been made to integrate human and environmental exposure assessment, with the most relevant being the EU FP7 project HEROIC (finished in October 2014; (Péry, Schüürmann et al. 2013, Wilks, Roth et al. 2015, Ciffroy, Péry et al. 2016)). Following an Expert Workshop on Extrapolations in Integrated Exposure Assessment that was organized in January 2014 as part of the HEROIC project it was concluded that the key points for integration across the human and environmental disciplines is the move from environmental fate and exposure estimations to the internal dose in the exposure assessment, which would provide a link between exposure and effects for a more relevant extrapolation between species, levels of biological organization and field monitoring. This move could be facilitated by the common use and sharing of emission and exposure data, i.e. distribution, fate and exposure models, monitoring data on concentrations in environmental media and food, PBPK and PBTK models, dose–response models and assessment of the variability for the critical effect across and within species, and the development of a common model for exposure assessment.

The most comprehensive tool that integrates human and environmental exposure assessments is EUSES used for industrial chemicals and biocides. The exposure assessment in EUSES covers the whole life cycle of substances as well as their fate in all environmental compartments providing exposure estimation for assessing both the environment and man indirectly exposed via the environment. The exposure assessment aims at reasonable worst case results by applying unfavourable, but not

unrealistic standard exposure scenarios and, as much as possible, mean, median or typical parameter values. Although the models used to simulate the fate in the different media could be further refined and/or made more specific by taking the environmental variation in local conditions into account, all the essential elements necessary for an integrated approach have been covered. Another promising tool is the modelling tool MERLIN-Expo (http://merlin-expo.4funproject.eu/), which is currently under construction in the frame of the EU FP7 4FUN. This tool is based on a library of models simulating the fate of chemicals (organic substances and metals) in the main environmental systems and in the human body, including models for non-biological receptor media (river, lake, soil, air), biological media of concern for humans (cultivated crops, cow milk), wildlife biota (primary producers in rivers, invertebrates, fish) and of Physiologically-Based PharmacoKinetic (PBPK) models.

The national research council published in 2014 a report on comparative exposure assessment aiming to develop a framework for assessing potentially safer substitute chemicals in terms of human health and ecological risks (National Research Council 2014). More recently, an article has been published by Teeguarden, Tan et al. (2016) that went further and proposed an aggregated exposure pathway (AEP) framework, in analogy to the adverse outcome pathway (AOP) concept applied in toxicology, and that has the potential to structure integrated exposure assessment (Figure 7). Based on these initiatives and the already obtained results further integration in exposure assessment does not seem warranted at this stage.

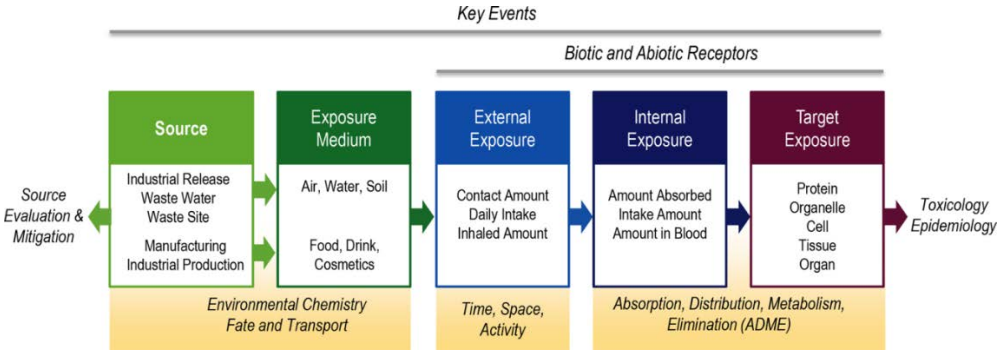


Figure 6 Aggregated exposure pathway (AEP) as proposed by Teeguarden, Tan et al. (2016). Principle components of an aggregate exposure pathway (AEP) cover all necessary levels of ecological, biological, and physical organization. Each box represents a key event which is a measurable change in a chemical state and concentration that is essential, but not necessarily sufficient, for the movement of a chemical from a source to the target site exposure. Each arrow represents a key event relationship which links a pair of key events. AEP's can be used to accumulate information for source mitigation, or use in epidemiology and toxicology.

5 Toxicokinetics & Toxicodynamics

Toxic effects are a manifestation of the internal dose or concentration at the target site, as well as the duration of exposure at that site (Vermeire, Munns et al. 2007). This process is captured by the toxicokinetics (TK) and toxicodynamics (TD) of a substance. The toxicokinetics describe the time-course of substance uptake, internal distribution, biotransformation and elimination in an organism, commonly abbreviated as ADME (Absorption, Distribution, Metabolism and Excretion). Together these processes determine the relationship between the external exposure concentration that is achieved in the HRA by substances application thorough swallowing, inhalation, injection, or skin contact, and in the ERA by substance spiking to media, soil, sediment and/or food, and the biologically effective dose, being the concentration at the organs and tissues where toxicity is exerted. The time course of toxic action at these target sites, the subsequent physiological impairment of the organism as well as the influence of any compensating mechanisms and finally the emergence of toxic effects at the level of the organism, are described by the toxicodynamics of a substance (Ashauer and Escher 2010).

5.1 TK & TD in human health risk assessment

In order to investigate the toxicokinetics of substances in mammals, the OECD has issued three Test Guidelines (TG): TG 417 'Toxicokinetics' as revised in 2010, TG 427 'Dermal absorption – *in vivo* method' and TG 428 'Dermal absorption –*in vitro* method' ((OECD 2004, OECD 2004, OECD 2010), respectively). Generally, these toxicokinetic studies are conducted separately from toxicity testing in order to obtain ADME/TK data. The application of TK data in the toxicity evaluation of substances is commonly applied during the drug development process, while in chemicals frameworks, such as, consumer products, food additives, biocides and industrial chemicals, this is less common (Creton, Saghir et al. 2012).

As highlighted by Bessems and Geraets (2013), information on ADME/TK acquired in an early phase of chemical safety assessment is useful to select a proper dosing regimen for toxicity testing (in case of non-linear kinetics occurring at high doses) or to select the most appropriate model species (in case one species turns out to be a clear outlier with respect to kinetics). Information on the internal exposure (i.e. AUC, C_{max}) would also be supportive to understand the observed health effects in a mode of action framework and to link the external dose to the internal exposure. In that respect determination of Margins of Safety between human exposure and animal points of departure (BMDL or NOAEL), should be based on AUC and C_{max} values in man and experimental animals.

Different authors have proposed an integrated approach where ADME/TK data is generated in parallel with toxicity studies (Saghir, Bartels et al. 2012, Bessems and Geraets 2013, Heringa, Brandon et al. 2013). Such an approach could help to avoid excessively high doses that

can cause unnecessary suffering in experimental animals, and could help to choose the most appropriate dosing regimen for the health effect testing (linear dose range; gavage or diet). As such, this information avoids irrelevant study design and thus unnecessary animal testing.

In the integrated approach proposed by Saghir, Bartels et al. (2012), generation of TK data without using additional animals can help to select the most appropriate animal species (for further testing) and to interpret results from initial toxicity testing and thus improve the quality of the risk characterisation (Figure 8). This approach includes determining pharmacokinetic (PK) behaviour of the test substance early in the testing program in a limited number of animals, and utilizing that information to determine *in silico* the optimal sampling times at steady-state exposure from dietary administration. The relationship between diurnal systemic and ingested doses is then determined for each dietary study and the highest doses for the subsequent longer-term studies are selected based either on the nonlinearity of the dose–response curve (kinetically derived-maximum dose and/or the conventional maximum-tolerated dose) of the preceding studies (Saghir, Bartels et al. 2012).

Bessems and Geraets (2013) proposed a testing strategy where the first step is to generate basic toxicokinetics data for a substance in order to steer (further) toxicity testing. Such TK data include information on species specificity (man versus rat, mouse, rabbit), exposure scenario versus toxicity testing scenario (e.g. vehicle, concentrations in matrix, skin surface area dose), route specificity (oral versus dermal versus inhalation) and/or bioavailability (sum of absorption and route-specific metabolism). They also propose that, instead of the generating ADME/TK data separately, ADME/TK investigations should be included in the regulatory requirements and as part of existing Test Guidelines for studying human health effects. This would prevent the risk that ADME/TK data are generated under different test conditions (route, dose, species/strain etc.) as compared to those of the critical health effect study. Furthermore, as TK and toxicity data are generated in the same animals, interpretation would likely be easier, and the number of test animals used could probably be reduced. Overall, this approach could help prevent unnecessary repeating of ADME/TK and health effect studies where another design (e.g. other species, other route) is to be followed.

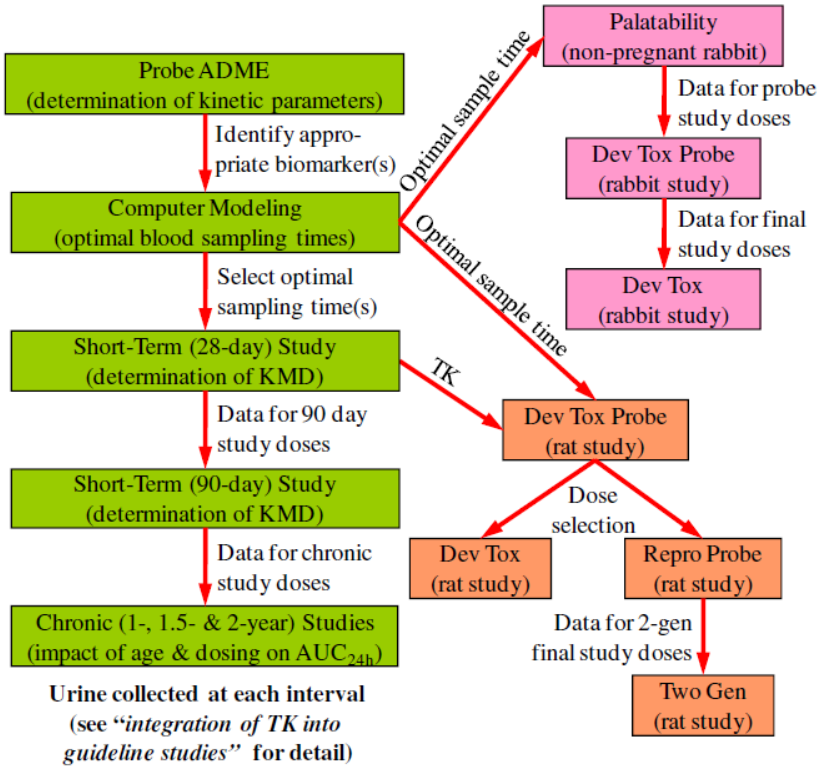


Figure 7 Study design for the integration of TK in toxicity testing of nonpharmaceuticals (obtained from Saghir, Bartels et al. (2012)).

Toxicokinetic data can also be used for exposure-based waiving, but are also necessary for the extrapolation of the safe concentrations determined *in vitro* to safe external doses *in vivo* for humans. Physiology-based pharmacokinetic (PBPK) models are a centerpiece in such an extrapolation. These need input of toxicokinetic data determined, for example, by *in vitro* kinetic tests or *in silico* predictions of kinetic parameter values. Some of the necessary toxicokinetic methods and tools for the new risk assessment strategy are already available, but others still need major further development (Heringa, Brandon et al. 2013, Sellers, Deleebeeck et al. 2015). There are not too many examples where PBTK models were used to extrapolate *in vitro* response to *in vivo* hazard assessment (Pery, Brochot et al. 2013). Rotroff, Wetmore et al. (2010) combined a simple PBTK model (one compartment only) with *in vitro* bioactivity data. Aim was to assess overlapping of chemicals between actual human oral exposure distributions and oral equivalent doses necessary to produce a steady-state *in vivo* concentration equivalent to *in vitro* concentration at 50% of maximum activity and lowest effective concentration values. Pery, Brochot et al. (2011) showed the possibility to relate *in vitro* and *in vivo* gene responses in macrophages following exposure to benzo(a)pyrene through PBTK modelling, while Punt, Schiffelers et al. (2011) showed that predictions of *in vivo* effects on the basis of integrated *in vitro* toxicity data and PBPK modelling approaches are generally within one order of magnitude of the observed *in vivo* situation. Pery, Brochot et al. (2013) demonstrated that *in vivo* hepatotoxicity can be predicted with an accuracy comparable to *in vivo* based methods using a framework

composed of a toxicodynamic model for cell viability and a PBPK model calibrated using *in silico*, *in vitro* or *in vivo* generated TK data.

5.2 TK & TD in environmental risk assessment

In the environmental risk assessment toxicokinetics is mostly investigated to determine the bioaccumulation potential of substances in aquatic and terrestrial organisms. The outcome is used in classification and labelling, the PBT assessment and the risk assessment for secondary poisoning. For this purpose, one-compartment models are commonly used to estimate a chemical concentration in the whole body of an organism, resulting in a bioconcentration factor (BCF) for aquatic organisms or a biota sediment/soil accumulation factor (BSAF) for sediment and terrestrial organisms. The BCF is the ratio of the concentration in an organism divided by the water concentration, where the water phase is the only exposure route. BCF values are mostly determined in the laboratory and can be determined at steady state or by kinetic fitting substance uptake and depuration as described in OECD TG 305 (OECD 2012). The BCF is expressed on a wet weight basis and preferably normalised to 5% lipids and corrected for growth (ECHA 2014). Within REACH (EC No. 1907/2006) and the Water Framework Directive (2000/60/EC), a BCF value greater than or equal to 100 is considered an indication of the potential for bioaccumulation, as it could impair the health of an organism or that of the organisms feeding upon it. Provided there are no mitigating substance properties such as rapid degradation (ready biodegradability or hydrolysis half-life <12h at pH 5-9, 20°C) or an obvious molecular size exclusion, a BCF ≥100 (or, if absent, a log K_{ow} ≥3) serves as a trigger for assessment for secondary poisoning of predators. Under the CLP regulation (EC No. 1272/2008), a BCF ≥500 (or, if absent, a log K_{ow} ≥4) is considered indicative of the potential to bioconcentrate for classification purposes. The CLP regulation furthermore states that some relationships can be observed between chronic toxicity and bioaccumulation potential, as toxicity is related to the body burden. Consequently, for substances that are not rapidly degradable and/or have a BCF≥500, in the absence of adequate chronic aquatic toxicity data, acute aquatic toxicity data can be used for classification for long-term aquatic hazards. Regarding the PBT assessment under REACH, a substance is considered to fulfil the bioaccumulation criterion (B) when the BCF in aquatic species is higher than 2000, and to fulfil the "very bioaccumulative" criterion (vB) when the BCF is higher than 5000.Models to derive a BCF value use octanol−water partition coefficients to quantify partitioning, assuming that the chemical is homogeneously circulated within the organism. Arnot and Gobas (2004) developed a model to predict the bioaccumulation processes of organic chemicals in aquatic ecosystems with the aim of providing data on site-specific toxicant concentrations and associated bioconcentration, bioaccumulation, and biota−sediment accumulation factors in organisms of aquatic food webs. Hendriks and Heikens (2001) developed a model that describes the accumulation kinetics of organic chemicals as a function of the octanol−water partition coefficient (K_{ow}), as well as the lipid content, weight, and trophic level of the species. For more information on BCF modelling and suitability for regulatory purposes, the reader is referred to the REACH guidance (ECHA 2014).

While hardly applied in regulatory context, considerable research has been conducted on the application of process-based toxicokinetic-toxicodynamic (TK-TD) models in environmental toxicology (e.g. Jager, Heugens et al. (2006) and Ashauer and Escher (2010)). The reader is referred to Ashauer and Escher (2010) for a comprehensive overview of the development of TK-TD modelling over time. Models discussed include; the time-to-event models (TTE) that provides information on the time needed until a certain event occurs (e.g., death) (Newman and McCloskey 1996); the Critical Body Residue (CBR) model that assumes that an organism dies when an internal threshold concentration (= Critical Body Residue or Lethal Body Burden) is exceeded; the Critical Target Occupation (CTO) model that assumes mortality occurs when a critical number of targets are irreversibly occupied (Legierse, Verhaar et al. 1999, Verhaar, de Wolf et al. 1999); the PULSETOX model that relates whole-body residues and acute lethality (Hickie, McCarty et al. 1995); the Damage Assessment Model (DAM) and Threshold Damage Model (TDM) that use damage to model toxicodynamics, and the DEBtox model that relates internal concentrations to sub-lethal effects based on the idea that toxicants modify the allocation of energy in an organism (Jager, Heugens et al. 2006, Billoir, Delignette-Muller et al. 2008). Ashauer and Escher (2010) also indicate that there is a need for TK-TD models for mechanisms of sub-lethal toxicity that do not act via modification of energy allocation. The main aim of these TK-TD models is to go beyond the descriptive character of summary statistics such as NOECs and EC_x, and to derive information from ecotoxicity tests that will allow a better understanding of the processes behind the effect. A combination of bioaccumulation kinetics and toxicity information can play an important role to formalize established knowledge about toxicity of compounds, sensitivity of organisms, organism recovery times and carry-over toxicity. TK-TD models might also be able to simulate temporal aspects of toxicity that can be used for risk assessment of fluctuating or pulsed exposures to pollutants.

5.3 Opportunities for integration of human health and environmental risk assessment with regard to toxicokinetics & toxicodynamics

In both the HRA and the ERA simple molecular properties of the compound, such as molecular size, lipophilicity, polar surface area (PSA) and hydrogen bonds are used to estimate ADME endpoints, such as the oral absorption, bioavailability, bioaccumulation, biliary excretion, blood brain barrier penetration and skin permeability (de la Nuez and Rodríguez 2008).

Generally, the larger a molecule, the larger a cavity must be formed in water in order to solubilise the compound, and thus the lower the water solubility. Increasing molecular size also reduces (oral) absorption, mainly by decreasing the compound's concentration at the surface of the intestinal epithelium, and by impairment of passive diffusion through the tightly packed aliphatic side chains of the bilayer membrane. Information on molecular size, expressed as either the length or the average maximum diameter (D_{max} aver) of a molecule, is also used as indicator for limited bioaccumulation potential of a substance in the PBT assessment. Under REACH a D_{max} aver greater than 17.4 Å has been set as threshold for hindered uptake due to molecular size (ECHA 2017).

The underlying thought being that the substance disturbs the entire interior structure of the lipid bilayer of cell membranes or is too bulky to pass through the cell membranes and therefore does not bioaccumulate to a significant amount.

Lipophilicity, typically quantified as the *n*-octanol-water partitioning coefficient log K_{ow} (see section 3.1.10), contributes to many substance properties, including solubility, permeability, potency, selectivity, and promiscuity. Generally, with increasing log K_{ow} the aqueous solubility of a substance decreases, which reduces its absorption in human skin. On the other hand, the uptake of organic substances in aquatic organisms is driven by their hydrophobicity. In the ERA, substances with log K_{ow} below 4.5 are considered to that an insufficient affinity for lipids to exceed the B criterion used in the PBT assessment, and further bioaccumulation testing in fish is not required (ECHA 2017). The exception being substances that either bind to proteins instead of lipids, or substances that have a high *n*-octanol-air partition coefficient (K_{oa}) where bioaccumulation can be more pronounced in air-breathing mammals compared to fish.

The polar surface area is used as a metric for the optimisation of a substance ability to permeate cells, and reflects the ability of solutes to leave hydrogen binding environments. Hydrogen bonds increase solubility in water and must be broken in order for the compound to permeate into and through the lipid bilayer membrane. Substances that have a large polar surface area (expressed in Å) tend to be poor at permeating cell membranes and have difficulties passing the blood brain barrier.

The relationships between molecular properties and ADME processes have been used to formulate rules of thumb, with the best known example being Lipinski's rule of five (RO5) that is commonly utilised to filter substances for oral absorption (Lipinski, Lombardo et al. 1997). It states that for a substance to be orally active it should meet at least three of the following four criteria: (1) molecular weight less than 500 daltons; (2) log K_{OW} not greater than 5; (3) no more than 5 hydrogen bond donors; (4) no more than 10 hydrogen bond acceptors. Such rules of thumb have many exceptions, and better predictions are made using (quantitative) structure-activity relationships ((Q)SARs)). QSARS are *in silico* models that have been developed to predict the ADME properties of a substance based on *in vivo* and *in vitro* data. *In vivo and in vitro* data are rather complex reflecting the multiple ADME processes that influence the pharmacokinetic characteristics of a substance. Therefore, the quality of computational approaches depends on the quality, complexity and understanding of the experimental data they are based on. QSARs do provide the opportunity to integrate HRA and ERA.

A more challenging area of potential cross-fertilization between HRA and ERA experts is the estimation of internal doses and subsequent concentrations at target sites. Considering that toxicokinetic studies are often not a standard requirement in HRA and are not frequently performed in the ERA, it is considered very informative when internal concentrations from one species could be extrapolated to another species, within or across taxa. This could support the estimation of

potential hazards in untested species, based on the assumption that the internal concentrations are indicative for acute toxicity and the toxicity at the target site. A concrete example worth exploring would be the relationship between a substance bioaccumulation potential in fish, its accumulation in rodents, and its half-life in human blood.

Physiologically-Based BioKinetic models (PBBK) could also be used for this purpose (section 5.1) in which substance-dependent parameters like partition coefficients between blood and organs, uptake, excretion or metabolism rates can be extrapolated from one species to another based on allometric factors (Campbell, Clewell et al. 2012). It should however be realised that difference in terms of elimination and metabolism could be substantial (Andersen, Clewell Iii et al. 2006, Connors, Du et al. 2013) and therefore extrapolations between species still have a high uncertainties. In the PBT assessment a high K_{ow} and K_{oa} value could also indicative for bioaccumulation in air-breathing mammals, which might require a toxicokinetic study to investigate this further. It may be worth investigating if in case the BCF value in fish is low, this concern could be waived assuming that a higher metabolism in mammals the bioaccumulation will be even lower. It would also be useful to further explore the model, like the one proposed by Hendriks, Traas et al. (2005), to estimate critical organism concentrations in algae, fish, and rats to estimate based on octanol-water partitioning, organism composition and a QSAR model and the databases on Km values in fish and humans that have been implemented in the QSAR Toolbox. The development of PBBK models has also increased the possibilities to provide a better interpretation of *in vitro* toxicity data for their relevance in terms of a toxic dose in the *in vivo* situation.

6 Toxicology

6.1 Human Health toxicity studies

Human Health risk assessment deals with a single species (human) and its goal is to protect individuals, while the goal of the environmental risk assessment is the protection of an ecosystem. The purpose of hazard identification is to evaluate the weight of evidence for adverse effects in humans based on assessment of all available data on toxicity and mode of action. It is designed to address primarily two questions: (a) whether an agent may pose a health hazard to humans, and (b) under what circumstances an identified hazard may be expressed. The hazard identification and dose-response assessment steps are primarily based on a number of different tests where mammalian animals are exposed to the chemical or test substance. Many toxicity tests examine specific types of *endpoints*, such as eye irritation, mutagenicity, reproduction, or cancer. Other tests are more general in nature, ranging from acute (single-exposure) studies to repeat dose (multiple-exposure) studies, in which animals are administered daily doses of a test substance.

6.1.1 HRA requirements in different regulatory frameworks

In the EU, the data requirement to assess the risk to human health for industrial chemicals depends on the tonnage produced with a minimum set of data at 10 tonnes (i.e. skin/ eye irritation, skin sensibilisation, mutagenicity and acute toxicity, repeated dose and reproductive/development toxicity screening), and the highest requirement for chemical produced >1000 tonnes (Annex I). As PPPs, biocides, human and veterinary drugs are used intentionally to give cause an adverse effect on target organisms, the human toxicological profile has to be explored in much more detail and therefore these high requirements are more standard. In case these substances have a specific mode of action, like for PPPs and biocides such as organophosphates or choline esterase inhibitors, additional neurotoxicity or immunotoxicity tests can be required. When relevant dermal adsorption might be needed to determine the systemic human exposure.

Since 2013 there is a prohibition in the EU to test finished cosmetic products and cosmetic ingredients on animals and therefore the safety can only be assessed by using alternative non-animal tests.

6.1.2 Human health in vivo toxicity tests

To investigate the chemical-induced effects to following in vivo toxicity test are most frequently performed

- Repeated Dose 28-Day Oral, dermal and inhalation Toxicity Study in Rodents, providing information on the possible health hazards likely to arise from repeated exposure (OECD 2008).
- Repeated Dose 90-Day Oral, dermal and inhalation Toxicity Study in Rodents, providing information on the possible health hazards likely to arise from repeated exposure over a prolonged period of time covering post-weaning maturation and growth well into adulthood (OECD 1998).

- Reproduction/Developmental Toxicity Screening Test, providing initial information on possible effects on reproduction and/or development, either at an early stage of assessing the toxicological properties of chemicals, or on chemicals of concern (OECD 2015).
- Combined Repeated Dose Toxicity Study with the Reproduction/Developmental Toxicity Screening Test, providing information on possible effects on male and female reproductive performance such as gonadal function, mating behaviour, conception, development of the conceptus and parturition, either at an early stage of assessing the toxicological properties of test chemicals, or on test chemicals of concern (OECD, 2015)
- Prenatal Developmental Toxicity Study, providing general information concerning the effects of prenatal exposure on the pregnant test animal and on the developing organism ; this may include assessment of maternal effects as well as death, structural abnormalities, or altered growth in the foetus (OECD 2001)
- Two-Generation Reproduction Toxicity Study, to provide general information concerning the effects of a test substance on the integrity and performance of the male and female reproductive systems (OECD 2001)
- Extended One-Generation Reproductive Toxicity Study, providing an evaluation of the pre- and postnatal effects of chemicals on development as well as a thorough evaluation of systemic toxicity in pregnant and lactating females and young and adult offspring

A detailed overview of mandatory and optional toxicological endpoints measured in these repeated dose toxicity tests is provided in Table 3.

6.1.3 Human health in vitro toxicity tests

For some endpoints, progress has been made in developing alternative *in vitro* test methods; OECD Test Guidelines using *in vitro* techniques are available for skin/eye corrosion and irritation, skin sensitisation, genotoxicity, developmental toxicity and endocrine disruption. In recent years, these alternative test methods have influenced regulatory decision-making, especially when coupled with *in silico* approaches and grouping of substances into chemical categories.

To assess skin irritants a variety of cell-based methods have been developed based on monolayer cultures of human and animal skin cells (keratinocytes), multi-layered (3-dimensional (3D)) cultures. Knowledge of the skin sensitisation pathway has also prompted the development of non-testing and testing approaches to estimate the response in different key events along the adverse outcome pathway: e.g. *In chemico* and *in vitro* assays to measure reactivity with proteins, In vitro assays to measure keratinocyte inflammatory responses or markers of dendritic / monocytic cell activation (OECD 2016). For several years bacterial and mammalian cell assays are used to determine the genetic toxicity of chemicals. Although there are no in vitro assays available to cover the broader range of potential adverse developmental toxicological outcomes, there are a few in vitro assays available that can be used as pre-screens for teratogenicity, such as assays measuring inhibition of differentiation or proliferation / growth of the mouse embryonic stem

cells (ES) , mouse 3T3 fibroblast cells , or micromass cultures of rat limb bud (INVITTOX protocol). Also whole rat embryo culture can be used to investigate delays in the development of certain organ systems or their malformation (https://eurl-ecvam.jrc.ec.europa.eu). Alternatively, the zebrafish has been a major model in biomedical research for several decades, resulting in an immense body of information on organisms development and molecular embryology (Nüsslein-Volhard 1994, Lele and Krone 1996, Dooley and Zon 2000, van der Ven, Schoonen et al. 2020). For investigating the endocrine disrupting properties of chemicals several *in vitro* methods have been developed and are presented in annex I as part of an example of a pathway based integration of human health and environmental testing and assessment.

In annex II a short overview is given of the *in vitro* methods that are currently used in risk assessment or classification and labelling. For *in silico* methods the reader is referred to the OECD guidance on grouping of chemicals (OECD 2014).

Table 3 Detailed overview of mandatory and optional toxicological endpoints measured in the most frequently performed repeated dose toxicity tests

Test	Mandatory endpoints	Optional endpoints
Toxicokinetics OECD TG 417	information on mass balance, absorption, bioavailability, tissue distribution, metabolism, excretion, and basic toxicokinetic parameters [e.g. AUC].	
Acute Oral Toxicity OECD TG 401, 420, 423, 425	mortality (LD50 establishment), body weight changes, clinical signs of toxicity, gross/histopathological changes	
Acute Dermal Toxicity OECD TG 402, 434	mortality, body weight changes, clinical signs of toxicity, skin irritation (local effects), gross/histopathological changes	
Acute Inhalation Toxicity OECD TG 403, 433, 436	mortality (LC50), non-lethal threshold concentration (LC01), clinical signs of toxicity, body weight changes, gross/histopathological changes	
Dermal Irritation/Corrosion OECD TG 404	irritation or corrosion to the skin: erythema and oedema, necrosis, ulcers, bleeding; discoloration, scars, alopecia, etc.	
Repeated Dose 28-Day Oral, dermal or inhalation Toxicity Study in Rodents (OECD TG 407, 410, 412)	Weight (Testes, Epididymides, Adrenals, Prostate + seminal vesicles with coagulating glands), Histopathology (Testes and Ovaries, Prostate + seminal vesicle with coagulating glands, Epididymides, Uterus, including cervix , adrenal, thyroid, Vagina)	weight (Ovaries, Uterus, including cervix, Thyroid), Histopathology (Vaginal smears, Male mammary glands, Pituitary), T3, T4, TSH
Repeated Dose 90-day Oral, dermal or inhalation Toxicity Study in Rodents (OECD TG 408, 409, 411, 413)	body weight and body weight changes; food and water consumption; toxic response data by sex and dose level, including signs of toxicity; nature, severity and duration of clinical observations (whether reversible or not); results of ophthalmological examination; sensory activity, grip strength and motor activity assessments (when available); haematological tests with relevant base-line values; clinical biochemistry tests with relevant base-line values; terminal body weight, organ weights and organ/body weight ratios; necropsy findings; histopathological findings	

Test	Mandatory endpoints	Optional endpoints
Reproduction/Developmental Toxicity Screening Test (OECD TG 421)	body weight/body weight changes; food consumption, and water consumption; fertility, gestation, gestation length; offspring, post-natal growth, clinical observations; oestrous cycle and cycle duration; number of live births and post-implantation loss; pup body weight; AGD; nipple retention in male pups, thyroid hormone levels	
Combined Repeated Dose Toxicity Study with the Reproduction/ Developmental Toxicity Screening Test (OECD TG 422)	body weight/body weight changes; food consumption, and water consumption; fertility, gestation, gestation length; offspring, post-natal growth, clinical observations; oestrous cycle and cycle duration; number of live births and post-implantation loss; pup body weight; AGD; nipple retention in male pups, thyroid hormone levels	
Neurotoxicity study in rodents OECD TG 424	Identification of adverse effects to the nervous system: characterisation of behavioural and/or neurological abnormalities using neuro-behavioural, neuropathological, neurochemical or electro-physiological examinations; determination of dose-and time-response relationships in order to estimate a NOAEL	
Developmental neurotoxicity study OECD TG 426	detection of gross neurologic and behavioural abnormalities, and evaluation of brain weights and neuropathology during postnatal development and adulthood.	
Prenatal Developmental Toxicity Study (OECD TG 414)	clinical sign; body weight, body weight change and gravid uterine weight; food/water consumption; uterine weight; number of corpora lutea; number of implantations, number and percent of live and dead foetuses and resorptions; number and percent of pre- and post-implantation losses; number and percent of live offspring; sex ratio; foetal body weight; external, soft tissue, and skeletal malformations and other relevant alterations; total number and percent of foetuses and litters with any external, soft tissue, or skeletal alteration, as well as the types and incidences of individual anomalies and other relevant alterations	body weight change corrected for gravid uterine weight
Two-Generation Reproduction Toxicity Study (OECD TG 416)	food and water consumption, body weight data for P and F1 animals selected for mating; litter and pup weight data; nature, severity and duration of clinical observations; time of death during the study or	Standardisation of litter sizes

Test	Mandatory endpoints	Optional endpoints
	whether animals survived to termination; toxic response data by sex and dose, including indices of mating, fertility, gestation, birth, viability, and lactation; toxic or other effects on reproduction, offspring, postnatal growth, etc.; necropsy findings; histopathological findings; number of P and F1 females cycling normally and cycle length; total cauda epididymal sperm number, percent progressively motile sperm, percent morphologically normal sperm, and percent of sperm with each identified abnormality; time-to-mating, including the number of days until mating; gestation length; number of implantations, corpora lutea, litter size, number of live births and post-implantation loss; number of pups with grossly visible abnormalities	
Extended One-Generation Reproductive Toxicity Study (OECD TG 443)	food and water consumption, body weight data for P and F1 animals selected for mating; litter and pup weight data; nature, severity and duration of clinical observations; time of death during the study or whether animals survived to termination; toxic response data by sex and dose, including indices of mating, fertility, gestation, birth, viability, and lactation; toxic or other effects on reproduction, offspring, postnatal growth, etc.; necropsy findings; histopathological findings; number of P and F1 females cycling normally and cycle length; total cauda epididymal sperm number, percent progressively motile sperm, percent morphologically normal sperm, and percent of sperm with each identified abnormality; time-to-mating, including the number of days until mating; gestation length; number of implantations, corpora lutea, litter size, number of live births and post-implantation loss; number of pups with grossly visible abnormalities, anogenital distance	F2, Cohort 2 and Cohort 3
Carcinogenicity Studies OECD TG 451	carcinogenic effects resulting from exposure to the active substance; full histopathology and establishment of organ specificity of tumours induced and the dose-response relationship; identification of the maximum dose eliciting no adverse effect	

6.2 Environmental toxicity studies

The aim of environmental toxicity risk assessment it not to protect individual organisms, but to protect the ecosystem. While this aim appears clear, specific protection goals have not been explicitly defined in EU legislation, where in article 3 of the consolidated EU Treaty (C:2016:202:TOC, date 7 June 2016) the following is stated: "*The Community shall have as its task, […] a high level of protection and improvement of the quality of the environment, […]*". This principle is reflected in article 191 on community policy for the environment where it is stated that the diversity of situations in the various regions should be taken into account, and in article 114 where it is repeated that all current product legislation aiming at harmonizing the internal market should refer to the principle of achieving the high level of protection in the requirements for an environmental risk assessment of chemicals or products (examples of current product legislation are Chemicals (REACH), Biocides, Veterinary Medicines, Medicines, Feed additives, Sludge, Fertilizers, Genetic Modified Organisms, Dangerous Substances).

For a robust and efficient environmental risk assessment, it is crucial that the questions *what* to protect, *where* to protect it and over *what time period*, should not be interpreted differently by risk assessors and risk managers under different regulations. In 2010, the EFSA panel on Plant Protection Products and their Residues (PPR Panel) published a scientific opinion on the development of specific protection goal options for environmental risk assessment of pesticides, and their residues (EFSA PPR Panel 2010). They did not consider in their analysis Directive 98/8/EC on Biocides and Regulation (EC) No 1907/2006 on industrial chemicals (REACH), as the general aims and processes proposed in these frameworks were considered similar to the plant protection products legislation. They did consider the following European legislation: the EU- Consolidated Version of the Treaty, the new Regulation (EC) No 1107/2009 on plant protection products, the Annexes to the Directive 91/414/EEC, the Sustainable Use of Pesticides Directive 2009/128/EC, the Directive 92/43/EEC (Habitat Directive), the Directive 2000/60/EC (Water Framework Directive), policy documents on soil protection, the Directive 2006/118/EC (the Groundwater Directive) and the Directive 2008/56/EC (Marine Strategy Framework Directive). Considering that PPP framework requires the highest level of environmental fate and effect information, far more than under REACH (see Annex III), it can be accepted that REACH and the on REACH based biocides frameworks were not considered.

The PPR panel presented a process for defining specific protection goal options, which uses the ecosystem services approach (Millennium Ecosystem Assessment 2005) as an overarching concept and which could be used in consultation processes with risk managers and stakeholders. It was proposed to define specific protection goals in 6 dimensions: biological entity, attribute, magnitude of effect, temporal and geographical scale of the effect, and the degree of certainty that the specified level of effect will not be exceeded. Specific protection goals were presented for 7 groups of organisms (microbes, algae, non-target plants (aquatic and terrestrial), aquatic invertebrates, terrestrial non-target arthropods including honeybees, terrestrial non-arthropod

invertebrates, and vertebrates), covering all ecosystem services which could potentially be affected by the use of pesticides. It was concluded that to ensure ecosystem services, taxa representative for the key drivers identified needed to be protected at the population level or higher. However, for aesthetic reasons (cultural ecosystem services) it could be decided to protect vertebrates at the individual level. To protect biodiversity, impacts would at least need to be assessed at the scale of the watershed/landscape. The PPR panel also emphasized the importance of a tiered approach for risk assessment, the essential linking of exposure and effect assessments in terms of spatial and temporal scales, and the relevance of ecological scenarios for appropriate pesticide risk assessments. The PPR panel further stressed that the presented concepts should be used as input for dialogues between risk managers and risk assessors.

6.2.1 ERA requirements in different regulatory frameworks

Environmental effect data requirements differ between frameworks, with the PPP framework requiring the most complete set, followed by biocides. The reason behind this is rather obvious, as both types of substances have been developed to be toxic to certain taxonomic groups, and as such pose a potentially higher risk to other non-target organisms when compared to substances that have not been developed with that specific purpose, e.g. industrial chemicals, pharmaceuticals, food and feed additives or cosmetics. Interestingly, for cosmetics there are no specific requirements for environmental effect data, but rather all available ecological and environmental effects data on the substance/compound/mixture should be reported and used for environmental risk assessment, even though some cosmetics components are used as preservatives to prevent bacterial growth, and could thus be considered as biocides. The data requirements of the different frameworks are shown in Annex III that was drafted as part of the EU-7 HEROIC project, and was adapted for this report. Requirements for nanomaterials have been removed, as for nanomaterials specific nanoform properties need to be reported, but the assessment is according to the relevant framework. In the table mammalian endpoints have been added as they are needed to assess secondary poisoning, which is conducted for industrial chemicals, biocides and plant protection products. Mammalian data are not generated just for the environmental risk assessment, instead the most sensitive relevant mammalian long-term toxicological endpoints (NOAEL) expressed as mg test compound/kg bw/day are derived from the human health part of the dossiers. For veterinary medicinal products with a high lipophilicity (= high log K_{ow}) secondary poisoning is assessed for the aquatic and the terrestrial compartment. In this framework, the first step is to conduct a worst-case estimate using QSAR estimated BCF values that assume no metabolism, and only if a risk is anticipated are experimental fish and earthworm bioconcentration data generated.

In all frameworks, triggers and tiers are applied. While the approaches can differ in setup, they do have in common that besides a base set, additional data is only requested if needed. Under REACH data requirements increase with increasing tonnage, while for human and veterinary medicinal products the prescribed dose indirectly determines if environmental fate and effect data need to be generated at all. Next to

these triggers, there are more general triggers that are applied in all frameworks and that are based on the physicochemical properties of a substance, i.e. lipophilicity (expressed by log K_{ow}) to determine if bioconcentration data need to be generated, adsorption potential to determine if soil, sediment and/or groundwater toxicity testing is needed, ready biodegradability as an initial screening for further biodegradation testing, etc. Using the generated data, a conservative risk assessment is conducted in the first tier usually based on short-term standardized laboratory test with single species, which is extended with long-term toxicity data when required. If a potential risk is identified, refinement can be conducted by performing modelling approaches, e.g. Species Sensitivity Distributions (SSD) or a more realistic prediction of the environmental concentration, if data is available. While not commonly conducted for industrial chemicals or pharmaceuticals, higher tier toxicity tests, i.e. micro-/mesocosms, semi-field or field studies, can be conducted. The latter more realistic toxicity studies are commonly performed in the PPP framework. The above discussed approach with triggers and tiers is important to maintain an efficient and affordable ERA, but it hardly takes into account information on the mode of action of a substance. In fact, only for substances with a suspected endocrine disrupting mode of action are additional tests, e.g. fish full life cycle toxicity test, requested in the different frameworks that contain mechanistic endpoints. There is an exception though, and that is the PPP framework, where depending on the mode of action of a pesticide different requirements are formulated with regard to the base set of aquatic toxicity. For all pesticides a base set needs to be tested consisting of a green alga, daphnia and fish. For insecticides, an additional arthropod species (e.g. Chironomus) needs to be tested, while for herbicides an additional non-green alga needs to be tested in addition to the base set. These additional tests, however, do not provide mechanistic insight.

6.2.2 Short description of environmental toxicity tests.

The different tests given above are briefly described in the Table 4, with the focussing being on the endpoints that are determined. In environmental effect studies endpoints are generally determined in whole organisms on the level of survival, growth impairment and reproduction, and basically no mechanistic information is obtained, this in contrast to the toxicokinetic and toxicodynamic endpoints determined in human health effect studies. Exceptions are the studies that are performed to study endocrine disruption in fish and amphibians, and that include mechanistic endpoints. These studies are discussed in subsequent paragraphs in more detail.

Fish Short Term Reproduction Assay (OECD TG 229, version 2 October 2012)

This test guideline describes an *in vivo* screening assay for fish reproduction where sexually mature male and spawning female fish are held together and exposed to a chemical during a limited part of their life-cycle of 21 days (OECD 2012). The assay is run with three test chemical concentrations and the necessary controls, including a carrier control if necessary. Recommended species are the fathead minnow (all endpoints validated), the Japanese medaka (subset validated, i.e. vitellogenin and secondary sex characteristics) and the zebrafish (subset

validated, i.e. vitellogenin). For the fathead minnow and medaka, four replicate test vessels are used for each treatment level and control(s); whereas for zebrafish, two replicates are used. The number of fathead minnows, medaka, and zebrafish per replicates is 6, 6 and 10, respectively. During the conduct of the assay, the apical endpoint egg production is measured quantitatively daily in each test vessel. At termination of the 21-day exposure period, two biomarker endpoints, vitellogenin and secondary sexual characteristics, are measured in males and females separately, as indicators of endocrine activity of the test chemical. Gonads of both sexes are also preserved and histopathology is an optional endpoint for assessing the reproductive fitness of the test animals and to add to the weight of evidence of other endpoints.

21-day Fish Assay: A Short-Term Screening for Oestrogenic and Androgenic Activity, and Aromatase Inhibition (OECD TG 230, version 7 September 2009)
This test guideline describes an *in vivo* screening assay for certain EDCs where sexually mature male and spawning female fish are held together and exposed to a chemical during a limited part of their life-cycle of 21 days (OECD 2009). The protocol is similar to OECD TG 229. The assay is run with three test chemical concentrations and the necessary controls, including a carrier control if necessary. Recommended species are the fathead minnow, the Japanese medaka and the zebrafish. For the fathead minnow, four replicate test vessels are used for each treatment level and control(s). Different from the protocol of OECD TG 229, two replicates are used for medaka and zebrafish. The number of fathead minnows, medaka, and zebrafish per replicates is 6, 10 and 10, respectively. At termination of the 21-day exposure period, biomarker endpoint vitellogenin is measured in fathead minnow, Japanese medaka and zebrafish, whereas biomarker endpoint secondary sex characteristics are measured in fathead minnow and Japanese medaka only. This test include only one or two biomarker endpoints and no apical endpoints are included. These biomarker endpoint(s) are measured in males and females as indicators of oestrogenic, aromatase inhibition or androgenic activity of the test chemical.

Androgenised Female Stickleback Screen (OECD GD 148, version 18 August 2011)
This guidance document describes a method to detect (anti-)androgenic chemicals using androgenised females of the stickleback fish species (*Gasterosteus aculeatus*) exposed for 3 weeks to the chemical tested (OECD 2011). The protocol is a variant of OECD TG 230 and in principle similar to OECD TG 230, with two major differences: *1)* only female fish are used, and *2)* all groups except controls receive 5µg/L dihydro-testosterone (DHT), in addition to the test chemical. DHT is used in order to induce a fully controlled moderate level of the androgen regulated protein spiggin in the female stickleback kidney, to allow the detection of (anti)androgens. The assay is normally initiated with fish sampled from populations that are in spawning condition to facilitate selection of female fish. The assay is conducted using at least three test concentrations, as well as a water control, a solvent control, a DHT control where DHT alone is administered at 5 µg /L, and a test substance control where the test substance is administered alone at the highest concentration tested. Recommended species is the three-spined

stickleback (*Gasterosteus aculeatus*). Two vessels per treatment (replicates; each containing 5 female fish) are used. The measurement of biomarker endpoint spiggin serves for the detection of chemicals with (anti)androgenic mode of action.

Fish Sexual Development Test (OECD TG 234, version 28 July 2011)
This test guideline is in fact an extended version of the fish early life stage test (OECD TG 210). It assesses early life-stage effects and potential adverse consequences of putative EDCs on fish sexual development (OECD 2011). Fish are exposed, from newly fertilized eggs (120 per concentration) until the completion of sexual differentiation at about 60 days post hatch, to at least three concentrations with minimum of 4 replicates. Five test concentrations are recommended if the data are to be used for risk assessment. Recommended species are the Japanese medaka (*Oryzias latipes*), zebrafish (*Danio rerio*) and three spined stickleback (*Gasterosteus aculeatus*) that are validated and the fathead minnow (*Pimephales promelas*) that was partially validated. In the test guideline, only the first three species, except fathead minnow, have been recommended. It is stated, however, that the stickleback should not be used for risk assessment because the validation data available so far showed that in this species the alterations of phenotypic sex ratio by the test substances were uncommon. Apical endpoints like hatchability, time to hatch, growth, weight and sex ratio can be used for indicating adverse effects of chemical. Biomarker endpoints including vitellogenin (VTG) and sex ratio enable the test to indicate the mode of action of chemical.

Medaka Extended One Generation Reproduction Test (MEOGRT, draft OECD TG)
The test is started by exposing sexually mature males and females (at least 12 wpf) in breeding pairs for 3 weeks, during which the chemical is distributed in the organism of the parental generation (F0). Recommended speies is the Japanese medaka (*Oryzias latipes*). As near as possible to the first day of the fourth week, eggs are collected to start the F1 generation. During rearing of the F1 generation (a total of 14 wpf), hatchability and survival are assessed. In addition, fish are sampled at 9-10 wpf for developmental endpoints and spawning is assessed for three weeks from 12 through 14 wpf. An F2 generation is started after the third week of the reproduction assessment and reared until completion of hatching. Primary emphasis is given to potential adverse effects on population relevant parameters including survival, gross development, growth and reproduction. In order to provide mechanistic information is supplied on vitellogenin, phenotypic secondary sex characteristics (SSC), sex ratio, and gonad histology.

The Amphibian Metamorphosis Assay (AMA, OECD TG 231, version 7 September 2009)
The Amphibian Metamorphosis Assay (AMA) is a screening assay intended to empirically identify substances which may interfere with the normal function of the hypothalamic-pituitary-thyroid (HPT) axis (OECD 2009). The general experimental design entails exposing stage 51 *Xenopus laevis* tadpoles to a minimum of three different concentrations of a test chemical and a dilution water control for 21 days. There are four replicates of each test treatment. Larval density at test initiation is

20 tadpoles per test tank for all treatment groups. The observational endpoints are hind limb length, snout to vent length (SVL), developmental stage, wet weight, thyroid histology, and daily observations of mortality.

The Larval Amphibian Growth and Development Assay (LAGDA, OECD TG 241, version 28 July 2015)
This test guideline of the Larval Amphibian Growth and Development Assay (LAGDA) describes a toxicity test with Xenopus that considers growth and development from fertilization through the early juvenile period (OECD 2015). It is an assay (typically 16 weeks) that assesses early development, metamorphosis, survival, growth, and partial reproductive maturation. The general experimental design entails exposing *X.laevis* embryos at Nieuwkoop and Faber (NF) stage 8-10 (3) to a minimum of four different concentrations of test chemical (generally spaced at not less than half-logarithmic intervals) and control(s) until 10 weeks after the median time to NF stage 62 in the control, with one interim sub-sample at NF stage 62 (≤ 45 post fertilization; usually around 45 days (dpf). There are four replicates in each test concentration with eight replicates for the control. Endpoints evaluated during the course of the exposure (at the interim sub-sample and final sample at completion of the test) include those indicative of generalized toxicity: mortality, abnormal behavior, and growth determinations (length and weight), as well as endpoints designed to characterize specific endocrine toxicity modes of action targeting oestrogen-, androgen-, or thyroid-mediated physiological processes (i,e. thyroid histopathology, gonad and gonad duct histopathology, abnormal development, plasma vitellogenin (optional), and genotypic/phenotypic sex ratios). Over a duration of approximate 16 weeks, the assay requires a total number of 480 animals, i.e., *X. laevis* embryos, (or 640 embryos, if a solvent control is used) to ensure sufficient power of the test for the evaluation of population-relevant endpoints such as growth, development and reproductive maturation.

Table 4 Overview of the toxicity test used for environmental risk assessment, including the toxicological endpoints measured and the test species used

Type of study / route / OECD TG	**Toxicological endpoint**	**Species**
Avian Acute Oral Toxicity; OECD TG 223 (or US EPA OPPTS 850.2100)	LD50 values (mg/Kg bw), NOEL, lethal threshold dose, time course of response and recovery, gross pathological findings	preferably performed with quail species (Japanese quail *Coturnix coturnix* japonica or Bobwhite quail *Colinus virginianus*). Also mallard duck (*Anas platyrhynchas*) can be used.
Acute Oral Toxicity; OECD TG 420, 423, 425	body weight changes; clinical signs of toxicity; gross histopathological changes	rat (mice and dogs are also relevant)
Neurotoxicity study in rodents; OECD TG 424	Identification of adverse effects to the nervous system; characterisation of behavioural and/or neurological abnormalities using neurobehavioural, neuropathological, neurochemical or electrophysiological examinations; determination of dose-and time-response relationships in order to estimate a NOAEL	rat
Avian dietary toxicity test; OECD TG 205, (or US EPA OPPTS 850.2200)	LC50 values (mg/Kg food), NOEC (if possible), lowest lethal concentration (LLC), time course of response and recovery, gross pathological findings	preferably performed with quail species (Japanese quail *Coturnix coturnix* japonica or Bobwhite quail *Colinus virginianus*). Also mallard duck (*Anas platyrhynchas*) can be used.
Avian reproduction test; OECD TG 206, (or US EPA OPPTS 850.2300)	NOEC values (mg/Kg bw/day)	northern bobwhite quail (*Colinus virginianus*) or Japanese quail (*Coturnix coturnix)* and a wild waterflow species (preferably

Type of study / route / OECD TG	Toxicological endpoint	Species
		mallad duck, *Anas platyrhynchas*).
Fish, Acute Toxicity Test; OECD TG 203	LC50, details of observed effects	rainbow trout *(Oncorhynchus mykiss)* and a warm water fish species
Fish, Prolonged Toxicity Test: 14-day Study; OECD TG 204	effects on growth, threshold level for lethal effects (ECx values), NOEC values, details of observed effects	rainbow trout (*Oncorhynchus mykiss*)
Fish, Juvenile Growth Test; OECD TG 215	effects on growth, threshold level for lethal effects (Ecx values), NOEC values, details of observed effects	rainbow trout (*Oncorhynchus mykiss*)
Fish early life stage toxicity test (ELS test); OECD TG 210	NOEC; effects on development, growth and behaviour, details of observed effects on fish early life stages; EC10 and EC20 values (if possible)	rainbow trout and other recommended species in OECD guideline
Fish full life cycle toxicity test (FLC test); US EPA protocol OPPTS 850.1500	NOEC, details of effects on survival and behavior, effects on reproduction of the parental and the viability of the filial generation; EC10 and EC20 values (if possible)	
Bioconcentration of active substance in fish; OECD TG 305	steady - state bioconcentration factors (BCF), uptake rate constants and depuration rate constants, incomplete excretion, metabolites formed in fish and if available information on organ-specific accumulation.	(several recommended species in Annex 3 OECD guideline)
Daphnia sp., Acute Immobilisation Test; OECD 202 Mysid Acute Toxicicty Test; US EPA OPPTS 850.1035	EC50 (24 and 48 hrs for Daphnia and 48 and 96 for Mysid) for immobilization; higher concentration causing no immobilization (if possible)	Daphnia species (preferably *Daphnia magna*) and additionally in certain circumstances. Mysid shrimp (*Americamysis bahia)* or freshwater non-crustacean species, e.g. Chironomus spp. (see comments)

Type of study / route / OECD TG	Toxicological endpoint	Species
Daphnia reproduction and growth; OECD TG 211	NOEC, details of observed effects, LOEC, ECx (e.g. EC_{10}, EC_{20} or EC_{50}) if possible.	Daphnia species (preferably *Daphnia magna*)
Mysid chronic toxicity test; US EPA OPPTS 850.1350	LC50, details of observed effects, NOEC	Mysid shrimp (*Americamysis bahia*)
Oyster Bioconcentration; US EPA OPPTS 850.1710	BCF	Eastern oysters (*Crassostrea virginica*)
Sediment-Water Chironomid Toxicity Using Spiked Sediment; OECD TG 218 Sediment-Water Chironomid Toxicity Using Spiked Water; OECD TG 219	NOEC values, measure of effects on survival and development (including effects on emergence of adults), LOEC, ECx (e.g. EC10, EC20 or EC50) if possible.	Chironomus riparius
toxicity test on marine/brackish molluscs; AST-E724	EC50	*Mytilus adilus* or *Macoma baltica*
Freshwater Alga and Cyanobacteria, Growth Inhibition Test; OECD TG 201	ErC50, LOEC, NOEC values for algal growth rate	Green alga (e.g. *Pheudokirchneriella subcapitata*, synonym *Selenastrum capricornutum*) and if required (see comment) blue green alga *Anabaena flos-aquae* or a diatom e.g. *Navicula pelliculosa*.
Lemna sp. Growth Inhibition Test; OECD TG 221; ASTM E 1913-04	EC50, LOEC, NOEC values; details of observed effects must be reported.	Lemna sp.

Type of study / route / OECD TG	Toxicological endpoint	Species
Honeybees, Acute Oral Toxicity Test; OECD TG 213 (based on EPPO Guideline 170, mentioned in 91/414)	LD10, LD20 and LD50 values, NOEC values, details of observed subleathal effects (µg active substance / bee)	honebee (*Apis mellifera*)
Honeybees, Acute Contact Toxicity Test; OECD TG 214 (based on EPPO Guideline 170, mentioned in 91/414)	LD10, LD20 and LD50 values, NOEC values, details of observed subleathal effects (µg active substance / bee)	honebee (*Apis mellifera*)
Honeybee brood feeding test; ICPBR Method - according to Oomen et al., EPPO Bulletin, Volume 22, pp 613 to 616, 1992, or Schur et al. Bulletin of Insectology 56(1) 91-96 (2003)	LD10, LD20 and LD50 and NOEC values for adults bees and larvae, details of observed subleathal effects (µg active substance / bee)	honebee (*Apis mellifera*)
effects on non-target arthropods other than bees (spray formulations). The testing should be carried out in accordance with guidelines that satisfy at least the requirements for testing specified in the "SETAC Guidance document on regulatory testing and risk assessment procedures for plant protection products with non-target arthropods (ESCORT 2)" ref: Mead - Briggs et al	LR50, NOEC, sublethal effects	Cereal aphid parasitoid *Aphidius rhopalosiphi* (*Hymenoptera: Braconidae*) and the predatory mite *Typhlodromus pyri* (*Acari: Phytoseiidae*) (for test compounds that are applied as foliar sprays)

Type of study / route / OECD TG	Toxicological endpoint	Species
2000 (Aphidius rhopalisiphi); Blumel et al 2000 (*Typhlodromus pyri*)		
effects on non-target arthropods other than bees (insect growth regulators or PPPs with specific modes of action). The testing should be carried out in accordance with guidelines that satisfy at least the requirements for testing specified in the "SETAC Guidance document on regulatory testing and risk assessment procedures for plant protection products with non-target arthropods (ESCORT 2)" Relevant references: Blumer et al 2000; Schmuck et al 2000; Bakker et al 2000; Vogt et al 2000	LR50, NOEC, sublethal effects	Typhlodromus pyri (Acari: Phytoseiidae) and one other species (e.g. *Coccinella septempunctata*, *Orius laevigatus* or *Chrysoperla carnea*)
Earthworm acute toxicity test; OECD TG 207 and ISO 11268-1:1993	LC50	*Eisenia foetida* or *Eisenia andrei*
Earthworm reproduction test (Eisenia foetida / Eisenia andrei); OECD 222	EC10, EC20, NOEC effects on growth, reproduction and behaviour	*Eisenia foetida* or *Eisenia andrei*

Type of study / route / OECD TG	Toxicological endpoint	Species
Inhibition of reproduction of Collembola by soil pollutants; ISO 11267 Predatory mite reproduction test in soil; OECD 226	EC10, EC20, NOEC effects on growth, reproduction and behaviour	*Folsomia candida* and *Hypoaspis aculeifer*
Soil Microorganisms: Nitrogen Transformation Test; OECD TG 216	ECx	
Soil Microorganisms: Carbon Transformation Test;OECD TG 217	ECx	
Terrestrial Plant Test: Seedling Emergence and Seedling Growth Test; OECD TG 208 Terrestrial Plant Test: Vegetative Vigour Test; OECD TG 227	EC_{50}, $ER_{50,}$ NOEC, LOEC values	Dose response test on at least 6 (to 10) species representing as many taxonomic groups as possible.
Activated Sludge, Respiration Inhibition Test (Carbon and Ammonium Oxidation); OECD TG 209	EC50, NOEC	

6.2.3 *Alternatives in environmental risk assessment*

So far only one *in vitro* toxicity test is available for environmental toxicity assessment, i.e. the Fish Embryo Acute Toxicity (FET) Test according to OECD TG 236 that has been adopted in 2013 (OECD 2013). In this test, newly fertilised zebrafish eggs are exposed to the test chemical for a period of 96 hrs, with the reported endpoint lethality, i.e. a LC50 is reported. While four apical observations are recorded as indicators of lethality: (i) coagulation of fertilised eggs, (ii) lack of somite formation, (iii) lack of detachment of the tail-bud from the yolk sac, and (iv) lack of heartbeat, no mechanistic information is obtained for further development of the AOP concept. There are, however, efforts directed to the integration of mechanistic endpoints in the zebrafish FET (ZFET), e.g. Piersma (2011) investigated the relative sensitivity of morphologically observed effects versus gene expression changes using an in-house developed embryonic development scoring system and transcriptomics. Other alternative methods applied in environmental toxicity testing, are QSAR estimations (e.g. ECOSAR).

6.2.4 *Developments in environmental risk assessment*

It is expected that increased focus on adverse outcome pathways (AOP), with a biochemical mechanistic understanding, will outdate nowadays endpoints in hazard assessment. An AOP is "a conceptual construct that portrays existing knowledge concerning the linkage between a direct molecular initiating event (e.g., a molecular interaction between a xenobiotic and a specific biomolecule) and an adverse outcome at a biological level of organization relevant to risk assessment" (Ankley, Bennett et al. 2010). Adverse outcome pathways provide a consistent structure and terminology for organizing toxicological or ecotoxicological information obtained at different levels of biological or ecological organization. Usual endpoints can be seen as the convergence of many AOPs. For instance, production of VTG can be inhibited by antagonism of the estrogen receptor, direct inhibition of 17β-estradiol (E2) synthesis, or indirect reduction of E2 production by feedback inhibition of steroidogenesis (Ankley, Bennett et al. 2010). Only endpoints related to classification and labelling, like reprotoxicity of cancerogenicity, would remain. Gold standard tests (on whole organisms) that are performed nowadays will be single points in overall pathways. In ecotoxicology, the focus will surely be on populations and communities. Thus, empirical data generated in laboratory testing should not only meet the requirements of mechanistic toxicology but also of population modeling to provide relevant information for ecological risk assessment (Kramer, Etterson et al. 2011).

6.3 Overview of similarities and differences in toxicological and ecotoxicological tests

As shown in Table 3, most assays used to assess the effect on human health (including acute and repeated dose toxicity tests, irritation/corrosion tests, sensitization tests, developmental and reproductive toxicity tests, endocrine disrupting tests, mutagenicity tests, carcinogenesis tests, and neurotoxicity tests) are performed in a few mammalian species e.g. rats and mice. In contrast, current ecotoxicological tests include acute toxicity tests in algae/aquatic plants, daphnia, fish and bacteria; acute and chronic toxicity tests on

development, growth, reproduction and endocrine effects in molluscs, daphnia, fish, and xenopus (see Table 4). Test guidelines on irritation/corrosion, sensitization, mutagenicity, carcinogenesis, and neurotoxicity are only available for toxicological studies but not for ecotoxicological studies. An overview of similarities and differences in toxicological and aquatic ecotoxicological tests is given in Table 5.

6.3.1 Toxicological and ecotoxicological similarities and differences in acute toxicity tests

Different from rats in the acute toxicity tests that are dosed once, organisms in current ecotoxicity tests for both short term and long term are continuously exposed to the chemicals tested. In mammalian toxicity test, test animals can be exposed via different exposure routes (e.g. air, food, gavage, dermal), whereas in ecotoxicity test organisms are exposed via the medium in which they live (e.g. water, soil and sediment) or the food they ingest.

Only a few mammalian species, e.g. rodents, are available for the acute toxicological tests; in contrast, for the environmental risk assessment species from different taxa including bacteria, cyanobacteria, algae/aquatic plants, crustacean, and fish are used. For fish, for example, the test guidelines recommends 7 different species including both warm water and cold water species. Furthermore, next to these aquatic species, terrestrial and sediment inhabiting species can also be tested if there is a concern for the respective compartments, including different types of worms, springtails, non-biting midges, terrestrial plants. These latter studies are not further discussed in this section though, as the same rational applies as for aquatic toxicity tests. Overall, due to different routes of exposure, the difference in organisms, the responses to a chemical exposure may vary greatly between toxicological and ecotoxicological test results.

6.3.2 Toxicological and ecotoxicological similarities and differences in repeated dose toxicity tests

Repeated dose toxicity tests in mammals are intended to investigate effects on a very broad variety of potential targets of toxicity. They provide information on the possible health hazards likely to arise from repeated exposure over a relatively limited period of time, and a prolonged period of time covering e.g. post-weaning maturation and growth well into adulthood, including the major toxic effects, indicate target organs and the possibility of accumulation, and can provide an estimate of a no-observed-adverse-effect level (NOAEL) of exposure which can be used in selecting dose levels for chronic studies and for establishing safety criteria for human exposure. They can also be used to study effects on the nervous, immune and endocrine systems. Regarding these particular endpoints, the repeated dose toxicity tests should identify chemicals with neurotoxic potential, which may warrant further in-depth investigation of this aspect, as well as chemicals that interfere with thyroid physiology. It may also provide data on chemicals that affect the male and/or female reproductive organs in young adult animals and may give an indication of immunological effects. Endpoints of repeated dose toxicity tests include body weight/body weight changes; food consumption, and water consumption; toxic response data by sex and dose level, including signs of toxicity; nature, severity

and duration of clinical observations; sensory activity, grip strength and motor activity assessments; haematological tests with relevant base-line values; clinical biochemistry tests with relevant base-line values; body weight at euthanasia and organ weight data; necropsy findings; a detailed description of all histopathological findings in different organs; absorption data if available; etc.

In comparison with toxicological repeated dose tests, less endpoints are used in ecotoxicological tests because effects on target organs are generally not considered to be population relevant. Focus is on impairment of development, growth, sexual maturation and reproduction. There are ecotoxicological test guidelines (e.g. OECD TG 215) where endpoints, such as histological observations on target organs, are included, however, these endpoints are often optional and cannot be used to derive a valid no observed effect concentration (NOEC). In addition to the fish juvenile growth test, the other toxicity tests listed in Table 5 may be comparable to the toxicological repeated dose tests, but are considered more relevant to developmental and reproductive toxicological tests. Details can be found in the relevant sections.

6.3.3 Toxicological and ecotoxicological similarities and differences in developmental toxicity tests

Developmental toxicity testing in mammals has been designed to provide general information concerning the effects of prenatal exposure on the pregnant test animal and on the developing organism; this may include assessment of maternal effects as well as death, structural abnormalities, or altered growth in the foetus. Functional deficits are not a part of the prenatal developmental toxicity study (OECD TG 414), but can be tested as adjunct or in a separate study. Information on testing for functional deficiencies and other postnatal effects can be found in the guidelines for the two-generation reproductive toxicity study (OECD TG 416) and the developmental neurotoxicity study (OECD TG 426). Similar to other repeated dose toxicity tests, the prenatal developmental toxicity guideline includes many endpoints on maternal toxic response (e.g. survival, pregnancy, abortion, early delivery; abnormal clinical sign, change in body weight and gravid uterine weight, food and water consumption; necropsy findings, including uterine weight). Developmental endpoints by dose for litters with implants (number of corpora lutea; number of implantations, number and percent of live and dead foetuses and resorptions; number and percent of pre- and post-implantation losses) and developmental endpoints by dose for litters with live foetuses (number and percent of live offspring; sex ratio; foetal body weight, preferably by sex and with sexes combined; external, soft tissue, and skeletal malformations and other relevant alterations; criteria for categorisation if appropriate; total number and percent of foetuses and litters with any external, soft tissue, or skeletal alteration, as well as the types and incidences of individual anomalies and other relevant alterations), are included (OECD 2001).
In ecotoxicology, developmental toxicity tests have been developed for two taxa, i.e. fish and amphibians. For fishes, several species are recommended, e.g in the early life stage toxicity test (OECD TG 210) six fish species have been specified. In this test, population relevant endpoints including hatchability, survival, and growth (body weight and

body length) are always considered. To include sexual development in fish, the early life stage toxicity test can be prolonged and the endpoint sex ratio can be added. In contrast, for amphibians only the African clawed frog (*Xenopus laevis*) is recommended. The two available amphibian test guidelines specify apical endpoints, but also biomarkers that indicate activity on the thyroid pathway and/or on the estrogen and androgen pathways. Endocrine endpoints include mortality and abnormalities, time to NF stage 62; histological changes (thyroid gland, liver, gonads and kidney), growth (length and weight), sex ratio, and vitellogenin.

6.3.4 Toxicological and ecotoxicological similarities and differences in reproductive toxicity tests

Reproduction toxicity testing in mammals has been designed to provide general information concerning the effects of a test substance on the integrity and performance of the male and female reproductive systems, including gonadal function, the oestrus cycle, mating behaviour, conception, gestation, parturition, lactation, and weaning, and the growth and development of the offspring. Reproduction toxicity tests can also provide information about the effects of the test substance on neonatal morbidity, mortality, and preliminary data on prenatal and postnatal developmental toxicity and serve as a guide for subsequent tests. Similar to repeated dose toxicity tests and developmental toxicity tests, many endpoints are included in the reproductive toxicity tests of mammals. These endpoints can be found in the relevant test guidelines.

Ecotoxicological reproductive toxicity tests have been developed for three taxa, i.e. molluscs, daphnia/crustaceans, and fish. As is the case for the developmental tests, several fish species are recommended. Fecundity, i.e. the number of eggs, is the main focus of all these ecotoxicological reproduction tests. In line with the reproduction toxicity tests for mammals, in the daphnia and fish test focus is both on the adults and the offspring, which is in contrast to molluscs where the focus is merely on the fecundity of the adults. Another difference between fish and the other taxa is that in addition to apical endpoints, the fish tests also includes biomarkers for endocrine EAS activity.

6.3.5 Toxicological and ecotoxicological similarities and differences in endocrine activity tests

There are both *in vitro* and *in vivo* OECD test guidelines for detecting EAS activity. These tests are specifically developed and are not considered as tests for detecting adverse effects. Normally, the endpoints in these *in vitro* and *in vivo* tests are limited.
In contrast to the toxicological tests for endocrine activities, only *in vivo* tests are available and the majority of tests can be used for detecting both adverse effects and endocrine activities.

6.4 Opportunities for integration of human health and environmental risk assessment with regard to toxicology

In terms of toxicology, the integration of ERA and HRA could be improved by increasing our mechanistic understanding of the toxicity exerted by a substance, preferably in different species that cover the range of organisms used for ERA and HRA. As explained in sections 6.1

and 6.2, the hazard characterization step is often associated with extensive in vivo toxicity testing using standardized guidelines or protocols. The resources associated with the generation of such hazard data, both in terms of time, cost and animal use, are significant, and the amount of mechanistic information obtained is rather limited considering that the current in vivo testing approach is primarily based on apical endpoints. This is especially the case for environmental toxicity tests where endpoints are determined in whole organisms on the level of survival, growth impairment and reproduction, with the exception being several fish and amphibian toxicity tests that focus on endocrine disruption and where endpoints at lower levels of biological organization are also included (see sections 6.2.2.1 to 6.2.2.7). Overall, the inclusion of toxicokinetic and toxicodynamic endpoints in environmental and human health toxicity studies would yield mechanistic information that is required to take the integration of ERA and HRA in terms of toxicology a step further. This is in agreement with the conclusion drawn by Wilks, Roth et al. (2015) as part of the Heroic project (EU FP7), who stated that ideally regulatory testing requirements should be revised so that resulting data provide some evidence on the MoA/AOP. They furthermore recommend that research should focus on commonalities in MoA/AOP across species, in particular across mammalian and non-mammalian animals (vertebrates in the first place) and their usability for hazard and risk extrapolations to humans.

Table 5 Overview of similarities and differences in toxicological and aquatic ecotoxicological tests

Category	**Toxicological tests**		**Aquatic toxicity tests**	
Acute toxicity tests				
	TG 401	acute oral toxicity test (rat)	TG 201	Freshwater alga and cyanobacteria growth inhibition test
	TG 402	acute dermal toxicity test (rat)	TG 202	Acute Daphina Immobilisation test
	TG 403	acute inhalation toxicity test (rat)	TG 203	Acute fish toxicity test
			TG 204	Fish prolonged Toxicity test: 14 day study
			TG 236	Fish embryo acute toxicity test
			TG 221	Lemna growth inhibition test
			TG 209	Activated Sludge, Respiration inhibition test
Irritation/Corrosion				
	TG 404	Acute Dermal Irritation/Corrosion	no comparable studies	
	TG 405	Acute Eye Irritation/Corrosion	no comparable studies	
Sensitisation				
	TG 406	Skin sensitisation	no comparable studies	
Repeated dose toxicity				
	TG 407	Repeated dose 28 day oral toxicity study in rodents	TG 215	Fish juvenile growth test
	TG 408	Repeated dose 90 day oral toxicity study in rodents		
	TG 409	Repeated dose 90 day oral toxicity study in non-rodents		
	TG 410	Repeated dose dermal toxicity 21/28 day study		
	TG 411	Subchronic dermal toxicity 90 day study		
	TG 412	Subacute inhalation toxicity 28-day study		
	TG 413	Subchronic inhalation toxicity 90 day study		
Developmental toxicity				
	TG 414	Prenatal Development Toxicity Study	TG 210	Fish early life stage toxicity test

Category	**Toxicological tests**		**Aquatic toxicity tests**	
			TG 212	Fish short term toxicity test on embryo and sac-fry stages
			TG 234	Fish sexual development test
			TG 231	Amphibian Metamorphosis Assay
			TG 241	Larval amphibian growth and development assay
Reproductive toxicity				
	TG 415	One-generation reproduction toxicity study	TG 229	Fish short term reproduction assay
			TG 211	Daphnia reproduction test
			TG 242	*Potamopyrgus antipodarum* Reproduction Test
			TG 243	*Lymnaea stagnalis* Reproduction Test
	TG 416	Two-generation reproduction toxicity	TG 240	Medaka extended one generation reproduction test
	TG 443	Extended One-generation reproductive toxicity study	TG 240	Medaka extended one generation reproduction test
	TG 421	Reproduction/Developmental Toxicity Screening Test	TG 229	Fish short term reproduction assay
	TG 422	Combined repeated dose toxicity study with reproduction/developmental toxicity screening test	TG 229	Fish short term reproduction assay
Endocrine activity				
	TG 440	Utreotrophic Bioassay in Rodents	TG 230	21-day fish assay
	TG 441	Hershberger Bioassay in Rats	TG 230	21-day fish assay
			GD 148	Sticklebeck assay
	TG 455, 456, 458 and 493	EAS *in vitro* assays	TG 230	21-day fish assay
Carcogenicity studies				
	TG 451	Carcinogenicity studies	no comparable studies	
	TG 452	Chronic toxicity studies	no comparable studies	

Category	**Toxicological tests**		**Aquatic toxicity tests**
	TG 453	Combined chronic toxicity/carcinogenicity studies	no comparable studies
Mutagenicity studies			
	TG 471-490	mutagenicity studies	no comparable studies
Toxicokinetics			
	TG 417	Toxicokinetics	Bioaccumulation assay
Neurotoxicity			
	TG 418-419	Delayed Neurotoxicity of Organophosphorus Sustances	no comparable studies
	TG 426	Developmental Neurotoxicity Study	no comparable studies

Adequate information on the mode or mechanism of action of a substance could facilitate the development and application of new in vitro assays, read-across approaches or inter-species extrapolation, and could, as noted by Tollefsen, Scholz et al. (2014), facilitate the initial hazard assessment of a substance. Such knowledge would also facilitate the further development of the AOP concept since adverse outcomes would be detected by observing apical endpoints, and could be linked to endpoints determined at lower levels of biological organization that are related to chemicals perturbation(s) of pathways at the molecular initiating event (MIE) and/or later key events. Studying these effects in environmental and human health model species, would further facilitate interspecies extrapolation through the identification of key events that are conserved across species (OECD 2011). As highlighted by Perkins, Ankley et al. (2013), numerous studies have identified conserved pathways for diseases in non-mammalian vertebrates and invertebrates. Nevertheless, caution is needed when extrapolating between species. For example, Brown, Gunnarsson et al. (2014) assessed the variability in the susceptibility of fish species to pharmaceuticals and concluded that while the drug targets can be conserved between species, evolutionary divergence in drug-target activation, physiology, behaviour and ecological life history make it difficult to predict population-level effects. The paper "Are adverse outcome pathways here to stay?" by Garcia-Reyero (2015) noted that there are indeed critical comments regarding the AOP concept, but the principle of a pathway based approach opens numerous opportunities to apply non-traditional approaches for understanding the risks of chemical exposure. It is further concluded that even if the specific adverse outcomes differ between species, the AOP concept provides an useful framework for extrapolating chemical effects across species, as the MIE and/or key events leading to adverse outcomes can be conserved in humans and other species. As exemplified by Perkins, Antczak et al. (2015) this extrapolation can happen at different levels of an AOP. For instance, at the MIE level, sequences or structures of proteins can be compared under the assumption that evolutionary conserved proteins may have conserved functions. At a pathway level, extrapolations might be more complex as one needs to consider the sequence of events and the dose or threshold required to activate these events. These pathways can be examined as discrete pathways or as networks using cross-species comparative genomics. The network approach is considered more holistic, but requires more knowledge and further development of the AOP framework. In a pathway context, alternative species including embryo tests can show similar effects to those found in mammals, although the concentration needed to have an effect and the potential mechanisms of compensation and recovery might be different (Garcia-Reyero 2015). The zebrafish embryo acute toxicity test (ZFET) seems particularly promising in this perspective, and as such could serve as a bridge between human and environmental effect assessment. Annex IV gives an example of a pathway based integration of human health and environmental testing and assessment of endocrine disrupting chemicals.

Integrating AOPs in a conceptual framework, will thus allow a better usage of data generated by in silico, in vitro and in vivo methods. For example, the robustness of read-across results from a tested chemical

to another untested chemical could be improved by showing that the chemicals are not only structurally similar but that they also act through the same AOP. AOPs can also inform the development of fixed testing strategies using several alternative methods in combination, and increase confidence in the integration of such information for prioritizing chemicals for further assessment (OECD 2011). For those substances where insufficient information is available, AOPs can be used to determine the type of additional information that is needed to confirm whether or not a chemical will cause a specific toxicological effect (Tollefsen, Scholz et al. 2014).Previously, the NoMiracle project (full: Novel Methods for Integrated Risk Assessment of Cumulative Stressors in Europe) concluded that the integration of human and ecological effect assessment could be improved by deriving and applying pathway and mode-of-action-based assessment factors based on a unified human and ecological classification scheme for toxicokinetic and –dynamic processes (Løkke 2010). It was also concluded that separation of uncertainty and variability in human and ecological assessment factors would stimulate the integration process because it would make the differences and similarities in both extrapolation procedures more transparent. Altogether, it is expected that integration of AOPs in a conceptual framework will benefit regulatory decision making, and that this can improve the integration of human and environmental effect assessment.

7 Potency

In toxicity testing, test organisms are ordinarily exposed to a range of test concentrations and the effect is observed, yielding a dose-response, i.e. effect-concentration, curve. Important aspects of the dose-response relationship are the threshold and potency. For many substances, no effect is observed below a certain dose, the threshold, as the exposed organism can deal with the substance, by processes such as metabolism or excretion (see chapter 5). At increasing concentrations, toxicity starts occurring and can either gradually build up or develop rapidly. This is evident from the slope of the concentration response curve, with the more potent the substance, the steeper the slope of the dose-response curve. The mechanism of action, being specific or non-specific, plays a role in this regard. However, the most important parameters determining potency are the ability of a substance to bind a target, i.e. its affinity, and once bound to a target to evoke a response, i.e. its efficacy.

For the effects assessment, traditionally, data from acute toxicity tests are analysed in a statistical manner to yield a certain percentage of effect/mortality (EC50/LC50, using regression analysis) or in the case of chronic toxicity to yield a No Observed Effect Concentration (NOEC, using hypothesis testing) or No Observed adverse Effect Level (NOAEL). The use of the NOEC/NOAEL for risk assessment purposes is not undisputed, and many authors have proposed to replace it (Kimmel and Gaylor 1988, Leisenring and Ryan 1992, Laskowski 1995, Crane and Newman 2000). The most important critique is that the NOEC/NOAEL itself does not give any information on the concentration-effect curve. Consequently, a NOEC/NOAEL does not mean that there is no effect, in extreme cases it has been shown that a NOEC can approach 100% (Crane and Newman 2000), while generally a NOEC corresponds to 10 to 34% effect. Gaylor (1992) and Allen, Kavlock et al. (1994) re-analysed large sets of developmental toxicity data obtained using a range of study designs, and showed that NOAELs on average corresponded to an effect of about 5 to 20%. The higher the variability in the data or the smaller the sample size, the higher the NOEC/NOAEL, thus in a way rewarding a suboptimal experimental setup. Janer, Hakkert et al. (2007) showed that NOAELs for the same subchronic study performed for the same chemical could differ a factor of 10 due to uncertainties associated with differences in strains used, endpoints selected, dosing regime and just chance. A clear drawback of the NOEC/NOAEL is that it is limited to one of the doses included in a study, which can lead to a factor two to four difference in NOEC/NOAEL just based on the chosen dosing regime. For example, OECD test guidelines recommend separation factors of 3.2 for chronic aquatic toxicity tests in algae (OECD 2011), daphnia (OECD 2012) and fish (OECD 2013), and two to four fold intervals in human health toxicity tests, such as two-generation reproduction study (OECD 2001) and the reproduction/ developmental toxicity screening test (OECD 2015). The NOEC/NOAEL does thus not represent a true biological threshold, and as such cannot directly be used to establish the lowest exposure level without risk. To

account for the uncertainties and limitations of the NOEC/NOAEL additional safety factors are used in the risk assessment.

Alternatively, using the concentration-effect relationship, a concentration can be estimated which corresponds to a specified degree of an adverse effect. For toxicity studies used in the human health risk assessment this concept is utilised in the benchmark dose (BMD) approach that was initially proposed by Crump (1984). This approach is considered a scientifically more advanced method compared to the NOAEL approach for deriving a point of departure/reference point (Hardy, Benford et al. 2017). The BMD approach uses the available dose-response information by fitting mathematical models to the data, i.e. biologically-based or curve-fitting models for response data in the range of empirical observation, and extrapolation by modelling below the range of observation (or if insufficient data are available by default procedures, e.g. linear, nonlinear, or both) (US EPA 2012). The EFSA Scientific Committee recommends to average the different models for calculating the BMD confidence interval (Hardy, Benford et al. 2017). The confidence interval of the benchmark dose estimate accounts for the statistical uncertainty in the data and the lower one-sided 95% confidence limit (BMDL) is used instead of the NOAEL. Selection of the benchmark response (BMR) level depends on several parameters, including the type of data (quantal vs. continuous), sensitivity of the study design and toxicity endpoint. For quantal data an extra risk of 10% compared to the point of departure is customarily used, but lower values, e.g. 5% for reproductive and developmental studies with nested study designs, can be used. For continuous data, the preferred approach is to determine the minimal level of change in an endpoint that is biologically significant, and define the BMR based on that amount of change. If it is not evident what level of response corresponds to an adverse effect, a change in the mean equal to one control standard deviation (or lower, e.g., half standard deviation for more severe effects) from the control mean can be used (US EPA 2012). To derive the BMDLs software packages are available with the best known being the benchmark dose software (BMDS) developed by the U.S. EPA (www.epa.gov/ncea), and the PROAST software developed by the RIVM (www.rivm.nl/proast), the latter also being suitable to derive effect concentrations for environmental risk assessment. Considering that the BMD approach allows the estimation of equipotent doses by interpolation between applied doses, it is considered a suitable tool for comparison of the potencies of different substances, or of the same substance under different exposure conditions (Hardy, Benford et al. 2017).

In the environmental risk assessment, the benchmark dose approach is not applied. The limitations associated with NOECs are recognized though, and in recently updated environmental toxicity OECD test guidelines the recommendation is included to either abandon the concept of the NOEC and replace it with regression based point estimates ECx (OECD 2011), or to report both a NOEC and an appropriate ECx (e.g. (OECD 2012, OECD 2013)). Generally, the critical effect concentration ECx corresponds to 10% effect, which is considered non-adverse. In environmental toxicology, efforts are made to develop methods that go beyond interpolating an x% effect level with an empirical statistical model as is the case in the ECx approach. The

DEBtox model is such a biologically-based method that departs from the processes behind the toxic effect, explicitly including time in the analysis (Kooijman 1996, Jager, Heugens et al. 2006). It also allows to use measurements of more endpoints from the same test, or tests with time-varying exposure (e.g. due to degradation).

Further, there are no opportunities for educated extrapolations to the population level, to time-varying exposure, or to other environmental conditions (e.g. other temperatures or limiting food levels). Jager, Heugens et al. (2006) therefore propose to use so-called biologically-based models, such as DEBtox, which uses all the results from all time points to estimate model parameters (such NOEC). The advantages are that the parameters can be estimated with greater accuracy and it can be used to estimate the ECx at any point in time even what will happen after exposure longer than the test duration.

8 Regulation

In the European Union several regulatory agencies are responsible for risk assessment of different categories of compounds according to their mandate from the European Commission. For (industrial) chemicals and biocides this is the European Chemicals Agency (ECHA), for plant protection products, food & feed additives the European Food Safety Authority (EFSA), for human and veterinary pharmaceuticals the European Medicals Agency (EMA), and for cosmetics the Scientific Committee on Consumer Safety (SCCS). Furthermore, the European Environmental Agency (EEA) provides assessments to supports the development of a sustainable environment. In the following sections, the main goals of the different regulations will be described with the focus on the data requirements and the way integrated testing strategies (including the use of alternative methods) can be implemented.

8.1 Industrial chemicals

Within the EU, The sectorial chemical legislation REACH (EC/1907/2006) establishes procedures for collecting and assessing information on the properties and hazards of industrial chemicals. Companies need to register their substances at European Chemicals Agency (ECHA), which requires information on intrinsic properties of a substance. Using these data the registrant(s) need to assess if their chemicals may cause adverse effects to human health and the environment. The standard information requirements are those which are required as a minimum to meet the registration obligations of REACH. They depend on the quantity of the substance that is manufactured or imported into the EU/EEA and are described in Annexes VI to X to REACH (for toxicology data requirements in HRA en ERA see Annex I and III, respectively). These minimum data requirements may be adapted as appropriate. This means that certain tests may be waived if justified that the test are technically not feasible, read-across can be applied, or validated QSARs can be applied. In order to reduce the number of tests on animals, REACH also promotes the use of alternative methods (e.g. *in vitro*, and *in silico*) for the hazard assessment of substances to be used in integrated testing strategies based on consideration of MoA and weight of evidence (WoE) schemes.

8.2 Biocides

The Biocidal Products Regulation (BPR, Regulation (EU) 528/2012) sets out the provisions for the placing biocidal products on the market. This regulation aims to improve the functioning of the biocidal products market in the EU, while ensuring a high level of protection for humans and the environment (ECHA website). The European Chemicals Agency (ECHA) is in charge of technical and scientific evaluation and coordination of all the applications for inclusion of active substances in Annex I (list of active substances with requirements agreed at the Community level for inclusion in biocidal products). ECHA also plays a key role in the centralized authorization of products.

The data requirements for biocidal active substances and biocidal products are set out in the Biocidal Products Directive and cover the following areas: Information on the applicant: (1) Information on identity, (2) Physical, chemical and technical properties, (3) Methods for identification and analysis, (4) Efficacy, (5) Toxicological studies and exposure, (6) Ecotoxicological studies and exposure, (6) Measures to protect people, animals and the environment. Under this new regulation it would be possible to waive requirements if data are not scientifically necessary or not relevant. To reduce costs and animal testing it also aims to render mandatory the sharing of toxicological studies. It also encourages the use of alternative testing methods. In Annex 2 of the regulation it is prescribed that "*Before new tests are carried out to determine the properties listed in this Annex, all available in vitro data, in vivo data, historical human data, data from valid (Q)SARs and data from structurally related substances (read-across approach) shall be assessed first"*. The use of the methods and their limitations in replacing *in vivo* test are discussed in guidance on the Biocidal products regulation published on the ECHA website.

8.3 Plant protection products (PPP)

The Plant Protection Products Regulation (Regulation (EC) No. 1107/2009) sets out the provisions for the placing of PPPs on the market of the European Community. The objective of the evaluation of pesticides residues in crops and animal products is to establish acceptable exposure based on good agricultural practice, which leads to the derivation of maximum residue levels (MRLs) for indirect exposure via the environment. The risk to the environment in general and non-target animals should also be assessed. The data requirements for pesticides used for the assessment are the most stringent amongst all the classes of chemicals and in general, pesticides are recognized as the most extensively tested. The toxicology data requirements are laid down in Commission Regulation (EU) No 283/2013 and 284/2013. An overview of the data requirement as presented in Annex I and III for HRA and ERA, respectively. In Article 62 of The Regulation (EC) No 1107/2009 the avoidance of unnecessary testing on vertebrate animals is prescribed, including the avoidance of duplicate testing, and for sharing of tests and studies involving vertebrate animals. When new OECD Test Guidelines become available that can replace *in vivo* tests, any new studies submitted for the purposes of plant protection regulation should be conducted using these alternatives. Paragraph 5 of the Annex, in both Regulations No 283/2013 and 284/2013, states that *in vivo* tests may only be used when no other validated (*in vitro* and *in silico*) methods are available.

Since PPP are supposed to act via a specific mode of action, the integration of human and environmental RAs are in particular meaningful for pesticides with common MoAs in both humans and wildlife species, even more when these PPP have some commonalities in the exposure assessment such as common sources and emissions, distribution routes and exposure scenarios for humans and non-target species (Wilks, Roth et al. 2015). This has been demonstrated by the RIVM and the US EPA in an example case study of a deterministic

integrated environmental and health RA for organophosphates in a typical farming community (Vermeire, McPhail et al. 2001).

8.4 Veterinary drugs

For the active ingredients of veterinary drugs the requirements and procedures for the marketing authorisation are laid down in Directive 2001/82/EC and in Regulation No 726/2004. The establishment of both an ADI and species-specific Maximum Residue limits (MRLs) should ensure the protection of consumers against possible harmful effects resulting from the exposure to residues of veterinary medicinal products present in foodstuffs. The marketing authorisation also requires the reporting of the potential risk for the environment. In its Annex, the Directive briefly the outlines of a two phase risk analysis concept bearing an exposure estimation and an ecotoxicity assessment. To assist this analysis and providing further guidance on the data requirement VICH guidelines has been developed and adopted by the Committee for Veterinary Medicinal Products (CVMP) of the EMA.

The Directive does not mention the possibility to use alternative methods. However, the VICH guidelines does recommend that these methods could be used when offering an equivalent assurance of safety as animal tests.

8.5 Cosmetics

Regulation (EC) N° 1223/2009 on cosmetic products is the main regulatory framework for finished cosmetic products when placed on the EU market, which requires cosmetics to cause no damage to human health. According to this regulation no cosmetic product may be tested on animals anywhere in the EU. Consequently, the safety of new cosmetics can only be tested with alternative methods.

Technical guidance for the risk assessment procedure is given in the Notes of guidance for testing of Cosmetic Ingredients for their Safety Evaluation. Based on this guidance the Scientific Committee on Consumer Safety provides the European Commission with scientific advice on the safety of non-food consumer products (e.g. cosmetic products and their ingredients, toys, textiles, clothing, personal care and household products) and services (e.g. tattooing, artificial sun tanning). The SCCS's advice is intended to enable risk managers to take the adequate and required actions in order to guarantee consumer protection. Besides default calculations of Margin of safety under use conditions, it is also concluded whether the use is safe or unsafe with respect to genotoxicity and carcinogenicity (Zweers and Vermeire 2007).

9 Conclusion

In this report the risk assessment components used in the human health and ecological risk assessment have been analysed and compared to determine for which components more synergy can be obtained to enhance the effectiveness and efficiency of risk assessment procedure for man and the environment.

With respect to substance identification, physicochemical and solvation data a high level of integration has already been achieved. Further integration is not expected to substantially improve RA or increase the overall efficiency.

Models to predict the emission and distribution of substances, are also rather comparable for both risk assessment frameworks, and integration to predict the exposure concentration in the environment for assess the risk for non-target species and human exposed via the environment is already in place. Further integration of the exposure assessment could be achieved by using internal instead of external dose, to facilitate extrapolation between species, levels of biological organization and field monitoring. This move could be achieved by better use of monitoring data on concentrations in environmental media and food, kinetic models (PBPK, PBTK and PBBK), dose–response models and assessment of the variability for the critical effect across and within species. In that respect it is be useful exploring the relationship between the bioaccumulation potential in invertebrates and vertebrate species. PBBK models could also provide possibility for the extrapolation from *in vitro* to *in vivo* situation.

At present the hazard assessment for the environment and human health are performed separately, mainly because human effect assessment aims at protecting individuals and ecological effect assessment aims at protection of the ecosystem, they concentrate on different adverse effects (individual effects versus population effects) and another level of biological organisation (individual organisms versus populations and ecosystems). Relatively little attention is paid to the similarities in the processes that govern the toxicity of a substance, mainly because the necessary mechanistic understanding was lacking. Due to recent advances in molecular and biological sciences our understanding of chemical perturbation(s) of adverse outcome pathways at key events is increasing. When sufficient information on AOPs that could be triggered by a chemical in question is available, toxicity data of different receptors can be exchanged if there is convincing evidence that the mechanisms involved are the same. General assessment factors can replaced by mode-of-action and pathway-related assessment factors. The species-barrier is expected to become less important in the extrapolation of toxicity data because of the increasing role of mechanistic information. The AOP concept will also provide a possibility to make better use of alternative methods.

References

Allen, B. C., R. J. Kavlock, C. A. Kimmel and E. M. Faustman (1994). "Dose-response assessment for developmental toxicity. II. Comparison of generic benchmark dose estimates with no observed adverse effect levels." Fundam Appl Toxicol **23**(4): 487-495.

Andersen, M. E., H. J. Clewell Iii, Y.-M. Tan, J. L. Butenhoff and G. W. Olsen (2006). "Pharmacokinetic modeling of saturable, renal resorption of perfluoroalkylacids in monkeys—Probing the determinants of long plasma half-lives." Toxicology **227**(1–2): 156-164.

Ankley, G. T., R. S. Bennett, R. J. Erickson, D. J. Hoff, M. W. Hornung, R. D. Johnson, D. R. Mount, J. W. Nichols, C. L. Russom, P. K. Schmieder, J. A. Serrrano, J. E. Tietge and D. L. Villeneuve (2010). "Adverse outcome pathways: a conceptual framework to support ecotoxicology research and risk assessment." Environ Toxicol Chem **29**(3): 730-741.

Arnot, J. A. and F. A. Gobas (2004). "A food web bioaccumulation model for organic chemicals in aquatic ecosystems." Environ Toxicol Chem **23**(10): 2343-2355.

Ashauer, R. and B. I. Escher (2010). "Advantages of toxicokinetic and toxicodynamic modelling in aquatic ecotoxicology and risk assessment." J Environ Monit **12**(11): 2056-2061.

Bessems, J. G. and L. Geraets (2013). "Proper knowledge on toxicokinetics improves human hazard testing and subsequent health risk characterisation. A case study approach." Regul Toxicol Pharmacol **67**(3): 325-334.

Billoir, E., M. L. Delignette-Muller, A. R. R. Péry and S. Charles (2008). "A Bayesian Approach to Analyzing Ecotoxicological Data." Environmental Science & Technology **42**(23): 8978-8984.

Brown, A. R., L. Gunnarsson, E. Kristiansson and C. R. Tyler (2014). "Assessing variation in the potential susceptibility of fish to pharmaceuticals, considering evolutionary differences in their physiology and ecology." Philosophical Transactions of the Royal Society B: Biological Sciences **369**(1656).

Burger, J. and M. Gochfeld (2001). "On developing bioindicators for human and ecological health." Environ Monit Assess **66**(1): 23-46.

Campbell, J. L., Jr., R. A. Clewell, P. R. Gentry, M. E. Andersen and H. J. Clewell, 3rd (2012). "Physiologically based pharmacokinetic/toxicokinetic modeling." Methods Mol Biol **929**: 439-499.

Ciffroy, P., A. R. R. Péry and N. Roth (2016). "Perspectives for integrating human and environmental exposure assessments." Science of The Total Environment **568**: 512-521.

Connors, K. A., B. Du, P. N. Fitzsimmons, A. D. Hoffman, C. K. Chambliss, J. W. Nichols and B. W. Brooks (2013). "Comparative pharmaceutical metabolism by rainbow trout (Oncorhynchus mykiss) liver S9 fractions." Environmental Toxicology and Chemistry **32**(8): 1810-1818.

Crane, M. and M. C. Newman (2000). "What level of effect is a no observed effect?" Environmental Toxicology and Chemistry **19**(2): 516-519.
Creton, S., S. A. Saghir, M. J. Bartels, R. Billington, J. S. Bus, W. Davies, M. P. Dent, G. M. Hawksworth, S. Parry and K. Z. Travis (2012). "Use of toxicokinetics to support chemical evaluation: Informing high dose selection and study interpretation." Regul Toxicol Pharmacol **62**(2): 241-247.
Crump, K. S. (1984). "A new method for determining allowable daily intakes." Fundam Appl Toxicol **4**(5): 854-871.
de la Nuez, A. and R. Rodríguez (2008). "Current methodology for the assessment of ADME-Tox properties on drug candidate molecules." Biotecnología Aplicada **25**: 97-110.
Dooley, K. and L. I. Zon (2000). "Zebrafish: a model system for the study of human disease." Curr Opin Genet Dev **10**(3): 252-256.
Dorne, J. L. and A. G. Renwick (2005). "The refinement of uncertainty/safety factors in risk assessment by the incorporation of data on toxicokinetic variability in humans." Toxicol Sci **86**(1): 20-26.
ECHA (2014). Guidance for identification and naming of substances under REACH and CLP. Version 1.4 - June 2016. Helsinki, Finland, European Chemicals Agency**:** 116.
ECHA (2014). Guidance on Information Requirements and Chemical Safety Assessment Chapter R.7c: Endpoint specific guidance. Version 2.0 - November 2014. Helsinki, Finland, European Chemicals Agency**:** 260.
ECHA (2014). Guidance on the Biocidal Products Regulation: Volume IV: Environment - Part A: Information Requirements. Version 1.1 – November 2014. Helsinki, Finland, European Chemicals Agency**:** 110.
ECHA (2016). Guidance on Information Requirements and Chemical Safety Assessment Chapter R.7a: Endpoint specific guidance. Version 5.0 - December 2016. Helsinki, Finland, European Chemicals Agency**:** 587.
ECHA (2016). Recommendations for nanomaterials applicable to Chapters R.7a and R.7c (Human Health Endpoints). Draft (Public) Version 2.0- May 2016. Helsinki, Finland, European Chemicals Agency**:** 22.
ECHA (2017). Guidance on Information Requirements and Chemical Safety Assessment Chapter R.11: PBT/vPvB assessment. Version 3.0 - June 2017. Helsinki, Finland, European Chemicals Agency**:** 158.
ECHA (2019). Options to address non-extractable residues in regulatory persistence assessment. Helsinki, Finland, European Chemicals Agency**:** 5.
EFSA PPR Panel (2010). "Scientific Opinion on the development of specific protection goal options for environmental risk assessment of pesticides, in particular in relation to the revision of the Guidance Documents on Aquatic and Terrestrial Ecotoxicology (SANCO/3268/2001 and SANCO/10329/2002)." EFSA Journal **8**(10).
EMEA (2004). Guideline on summary of requirements for active substances in the quality part of the dossier. London, UK, European Medicines Agency**:** 10.

EMEA (2007). Guideline on excipients in the dossier for application for marketing authorisation of a medicinal product. London, UK, European Medicines Agency**:** 12.
European Commission (2008). Council 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) (Text with EEA relevance). Brussels, Belgium, European Commission.
Ferguson, L. N. (1956). "Hydrogen bonding and physical properties of substances." Journal of Chemical Education **33**(6): 267.
Garcia-Reyero, N. (2015). "Are Adverse Outcome Pathways Here to Stay?" Environmental Science & Technology **49**(1): 3-9.
Gaylor, D. W. (1992). "Incidence of developmental defects at the no observed adverse effect level (NOAEL)." Regul Toxicol Pharmacol **15**(2 Pt 1): 151-160.
Hardy, A., D. Benford, T. Halldorsson, M. J. Jeger, K. H. Knutsen, S. More, A. Mortensen, H. Naegeli, H. Noteborn, C. Ockleford, A. Ricci, G. Rychen, V. Silano, R. Solecki, D. Turck, M. Aerts, L. Bodin, A. Davis, L. Edler, U. Gundert-Remy, S. Sand, W. Slob, B. Bottex, J. C. Abrahantes, D. C. Marques, G. Kass and J. R. Schlatter (2017). "Update: use of the benchmark dose approach in risk assessment." EFSA Journal **15**(1).
Hendriks, A. J. and A. Heikens (2001). "The power of size. 2. Rate constants and equilibrium ratios for accumulation of inorganic substances related to species weight." Environ Toxicol Chem **20**(7): 1421-1437.
Hendriks, A. J., T. P. Traas and M. A. J. Huijbregts (2005). "Critical Body Residues Linked to Octanol−Water Partitioning, Organism Composition, and LC50 QSARs: Meta-analysis and Model." Environmental Science & Technology **39**(9): 3226-3236.
Heringa, M. B., E. F. A. Brandon, J. G. Bessems and P. M. J. Bos (2013). Integration of toxicokinetics and toxicodynamics testing essential for risk assessment. Bilthoven, the Netherlands, National Institute for Public Health and the Environment (RIVM)**:** 227.
Hickie, B. E., L. S. McCarty and G. D. Dixon (1995). "A residue-based toxicokinetic model for pulse-exposure toxicity in aquatic systems." Environmental Toxicology and Chemistry **14**(12): 2187-2197.
Jager, T., E. H. W. Heugens and S. A. L. M. Kooijman (2006). "Making sense of ecotoxicological test results: Towards application of process-based models." Ecotoxicology **15**(3): 305-314.
Janer, G., B. C. Hakkert, A. H. Piersma, T. Vermeire and W. Slob (2007). "A retrospective analysis of the added value of the rat two-generation reproductive toxicity study versus the rat subchronic toxicity study." Reprod Toxicol **24**(1): 103-113.
JRC (2011). REACH Implementation Project Substance Identification of Nanomaterials (RIP oN1). Advisory Report., Joint Research Centre 76.
Kästner, M., S. Trapp and A. Schäffer (2018). Consultancy services to support ECHA in improving the interpretation of Non-Extractable Residues (NER) in degradation assessment. Helsinki, Finland**:** 76.
Kimmel, C. A. and D. W. Gaylor (1988). "Issues in qualitative and quantitative risk analysis for developmental toxicology." Risk Anal **8**(1): 15-20.

Kooijman, S. A. L. M. (1996). "An Alternative for NOEC Exists, but the Standard Model Has to Be Abandoned First." Oikos **75**(2): 310-316.
Kostal, J., A. Voutchkova-Kostal, P. T. Anastas and J. B. Zimmerman (2015). "Identifying and designing chemicals with minimal acute aquatic toxicity." Proc Natl Acad Sci U S A **112**(20): 6289-6294.
Kramer, V. J., M. A. Etterson, M. Hecker, C. A. Murphy, G. Roesijadi, D. J. Spade, J. A. Spromberg, M. Wang and G. T. Ankley (2011). "Adverse outcome pathways and ecological risk assessment: bridging to population-level effects." Environ Toxicol Chem **30**(1): 64-76.
Laskowski, R. (1995). "Some good reasons to ban the use of NOEC, LOEC and related concepts in ecotoxicology." OIKOS **73**: 140-144.
Legierse, K. C. H. M., H. J. M. Verhaar, W. H. J. Vaes, J. H. M. De Bruijn and J. L. M. Hermens (1999). "Analysis of the Time-Dependent Acute Aquatic Toxicity of Organophosphorus Pesticides: The Critical Target Occupation Model." Environmental Science & Technology **33**(6): 917-925.
Leisenring, W. and L. Ryan (1992). "Statistical properties of the NOAEL." Regul Toxicol Pharmacol **15**(2 Pt 1): 161-171.
Lele, Z. and P. H. Krone (1996). "The zebrafish as a model system in developmental, toxicological and transgenic research." Biotechnol Adv **14**(1): 57-72.
Lipinski, C. A., F. Lombardo, B. W. Dominy and P. J. Feeney (1997). "Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings." Advanced Drug Delivery Reviews(23): 3-25.
Little Pro. (2016). "CRA Basics Part 1: Physicochemical Properties." Retrieved 2016-07-01, from http://www.chemsafetypro.com/Topics/CRA/physio-chemical_properties.html.
Løkke, H. (2010). "Novel methods for integrated risk assessment of cumulative stressors — Results from the NoMiracle project." Science of The Total Environment **408**(18): 3719-3724.
Millennium Ecosystem Assessment (2005). Ecosystems and Human Well-being: Synthesis. Washington, DC, Island Press.
National Research Council (2014). A Framework to Guide Selection of Chemical Alternatives. Washington DC, National Academies Press.
Newman, M. C. and J. T. McCloskey (1996). "Time-to-event analyses of ecotoxicity data." Ecotoxicology **5**(3): 187-196.
Nüsslein-Volhard, C. (1994). "Of flies and fishes." Science **266**(5185): 572-574.
OECD (1981). Test No. 110: Particle Size Distribution/ Fibre Length and Diameter Distributions Paris, France, Organisation for Economic Co-operation and Development**:** 13.
OECD (1981). Test No. 112: Dissociation Constants in Water. Paris, France, Organisation for Economic Co-operation and Development**:** 7.
OECD (1995). Test No. 102: Melting Point/ Melting Range Paris, France, Organisation for Economic Co-operation and Development**:** 8.
OECD (1995). Test No. 103: Boiling Point. Paris, France, Organisation for Economic Co-operation and Development**:** 8.
OECD (1995). Test No. 105: Water Solubility. Paris, France, Organisation for Economic Co-operation and Development**:** 7.

OECD (1995). Test No. 107: Partition Coefficient (n-octanol/water): Shake Flask Method. Paris, France, Organisation for Economic Co-operation and Development: 4.
OECD (1995). Test No. 115: Surface Tension of Aqueous Solutions Paris, France, Organisation for Economic Co-operation and Development: 7.
OECD (1998). Test No. 408: Repeated Dose 90-Day Oral Toxicity Study in Rodents. Paris, France, Organisation for Economic Co-operation and Development: 10.
OECD (2001). Test No. 106: Adsorption -- Desorption Using a Batch Equilibrium Method. Paris, France, Organisation for Economic Co-operation and Development: 44.
OECD (2001). Test No. 121: Estimation of the Adsorption Coefficient (Koc) on Soil and on Sewage Sludge using High Performance Liquid Chromatography (HPLC). Paris, France, Organisation for Economic Co-operation and Development: 11.
OECD (2001). Test No. 414: Prenatal Development Toxicity Study. Paris, France, Organisation for Economic Co-operation and Development: 11.
OECD (2001). Test No. 416: Two-Generation Reproduction Toxicity. Paris, France, Organisation for Economic Co-operation and Development: 13.
OECD (2002). OECD series on testing and sssessment. Number 23. Guidance document on aquatic toxicity testing of difficult substances and mixtures. Paris, France, Organisation for Economic Co-operation and Development: 53.
OECD (2004). Test No. 117: Partition Coefficient (n-octanol/water), HPLC Method Paris, France, Organisation for Economic Co-operation and Development: 11.
OECD (2004). Test No. 312: Leaching in Soil Columns. Paris, France, Organisation for Economic Co-operation and Development: 15.
OECD (2004). Test No. 427: Skin Absorption: In Vivo Method. Paris, France, Organisation for Economic Co-operation and Development: 8.
OECD (2004). Test No. 428: Skin Absorption: In Vitro Method. Paris, France, Organisation for Economic Co-operation and Development: 8.
OECD (2006). Test No. 104: Vapour Pressure Paris, France, Organisation for Economic Co-operation and Development: 18.
OECD (2006). Test No. 123: Partition Coefficient (1-Octanol/Water): Slow-Stirring Paris, France, Organisation for Economic Co-operation and Development: 16.
OECD (2008). Test No. 407: Repeated Dose 28-day Oral Toxicity Study in Rodents. Paris, France, Organisation for Economic Co-operation and Development: 13.
OECD (2009). Test No. 230: 21-day Fish Assay: a short-term screening for oestrogenic and androgenic activity, and aromatase inhibition. Paris, France, Organisation for Economic Co-operation and Development: 38.
OECD (2009). Test No. 231: The amphibian Metamorphosis Assay. Paris, France, Organisation for Economic Co-operation and Development: 33.
OECD (2010). Test No. 417: Toxicokinetics. Paris, France, Organisation for Economic Co-operation and Development: 20.

OECD (2011). OECD series on testing and assessment. Number 138. Report of the Workshop on Using Mechanistic Information in Forming Chemical Categories. Paris, France, Organisation for Economic Co-operation and Development**:** 176.
OECD (2011). OECD series on testing and assessment. Number 148. Guidance document on the androgenised female stickleback screen. Paris, France, Organisation for Economic Co-operation and Development**:** 47.
OECD (2011). OECD series on testing and assessment. Number 215. Report of the workshop on a framework for the development and use of integrated approaches to testing and assessment. Paris, France, Organisation for Economic Co-operation and Development**:** 154.
OECD (2011). Test No. 201: Freshwater Alga and Cyanobacteria, Growth Inhibition Test. Paris, France, Organisation for Economic Co-operation and Development**:** 25.
OECD (2011). Test No. 234: Fish sexual development test. Paris, France, Organisation for Economic Co-operation and Development**:** 44.
OECD (2012). Test No. 114: Viscosity of Liquidsd. Paris, France, Organisation for Economic Co-operation and Development**:** 6.
OECD (2012). Test No. 211: Daphnia magna Reproduction Test. Paris, France, Organisation for Economic Co-operation and Development**:** 25.
OECD (2012). Test No. 229: Fish Short Term Reproduction Assay. Paris, France, Organisation for Economic Co-operation and Development**:** 40.
OECD (2012). Test No. 305: Bioaccumulation in Fish: Aqueous and Dietary Exposure. Paris, France, Organisation for Economic Co-operation and Development**:** 72.
OECD (2013). Test No. 122: Determination of pH, Acidity and Alkalinity Paris, France, Organisation for Economic Co-operation and Development**:** 6.
OECD (2013). Test No. 210: Fish, Early-life Stage Toxicity Test. Paris, France, Organisation for Economic Co-operation and Development**:** 24.
OECD (2013). Test No. 236: Fish Embryo Acute Toxicity (FET) Test. Paris, France, Organisation for Economic Co-operation and Development**:** 22.
OECD (2014). OECD series on testing and assessment. Number 194. Guidance on grouping of chemicals, second edition. Paris, France, Organisation for Economic Co-operation and Development**:** 141.
OECD (2015). Test No. 241: The Larval Amphibian Growth and Development Assay (LAGDA). Paris, France, Organisation for Economic Co-operation and Development**:** 38.
OECD (2015). Test No. 421: Reproduction/Developmental Toxicity Screening Test. Paris, France, Organisation for Economic Co-operation and Development**:** 14.
OECD (2016). OECD series on testing and assessment. Number 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. Paris, France, Organisation for Economic Co-operation and Development**:** 24.

OECD (2016). OECD Workshop Report. Series on the Safety of Manufactured Nanomaterials. Number 63. Physical-chemical parameters: measurements and methods relevant for the regulation of nanomaterials. Paris, France, Organisation for Economic Co-operation and Development**:** 36.

Perkins, E. J., G. T. Ankley, K. M. Crofton, N. Garcia-Reyero, C. A. LaLone, M. S. Johnson, J. E. Tietge and D. L. Villeneuve (2013). "Current perspectives on the use of alternative species in human health and ecological hazard assessments." Environ Health Perspect **121**(9): 1002-1010.

Perkins, E. J., P. Antczak, L. Burgoon, F. Falciani, N. Garcia-Reyero, S. Gutsell, G. Hodges, A. Kienzler, D. Knapen, M. McBride and C. Willett (2015). "Adverse Outcome Pathways for Regulatory Applications: Examination of Four Case Studies With Different Degrees of Completeness and Scientific Confidence." Toxicol Sci **148**(1): 14-25.

Pery, A. R., C. Brochot, S. Desmots, M. Boize, L. Sparfel and O. Fardel (2011). "Predicting in vivo gene expression in macrophages after exposure to benzo(a)pyrene based on in vitro assays and toxicokinetic/toxicodynamic models." Toxicol Lett **201**(1): 8-14.

Pery, A. R., C. Brochot, F. A. Zeman, E. Mombelli, S. Desmots, M. Pavan, E. Fioravanzo and J. M. Zaldivar (2013). "Prediction of dose-hepatotoxic response in humans based on toxicokinetic/toxicodynamic modeling with or without in vivo data: a case study with acetaminophen." Toxicol Lett **220**(1): 26-34.

Péry, A. R. R., G. Schüürmann, P. Ciffroy, M. Faust, T. Backhaus, L. Aicher, E. Mombelli, C. Tebby, M. T. D. Cronin, S. Tissot, S. Andres, J. M. Brignon, L. Frewer, S. Georgiou, K. Mattas, J. C. Vergnaud, W. Peijnenburg, E. Capri, A. Marchis and M. F. Wilks (2013). "Perspectives for integrating human and environmental risk assessment and synergies with socio-economic analysis." Science of The Total Environment **456–457**: 307-316.

Piersma, A. H. (2011). Development and implementation of alternative methods in reproductive toxicology. Bilthoven, the Netherlands, National Institute for Public Health and the Environment (RIVM)**:** 35.

Punt, A., M. J. Schiffelers, G. Jean Horbach, J. J. van de Sandt, G. M. Groothuis, I. M. Rietjens and B. J. Blaauboer (2011). "Evaluation of research activities and research needs to increase the impact and applicability of alternative testing strategies in risk assessment practice." Regul Toxicol Pharmacol **61**(1): 105-114.

Rotroff, D. M., B. A. Wetmore, D. J. Dix, S. S. Ferguson, H. J. Clewell, K. A. Houck, E. L. Lecluyse, M. E. Andersen, R. S. Judson, C. M. Smith, M. A. Sochaski, R. J. Kavlock, F. Boellmann, M. T. Martin, D. M. Reif, J. F. Wambaugh and R. S. Thomas (2010). "Incorporating human dosimetry and exposure into high-throughput in vitro toxicity screening." Toxicol Sci **117**(2): 348-358.

Saghir, S. A., M. J. Bartels, D. L. Rick, A. T. McCoy, R. J. Rasoulpour, R. G. Ellis-Hutchings, M. Sue Marty, C. Terry, J. P. Bailey, R. Billington and J. S. Bus (2012). "Assessment of diurnal systemic dose of agrochemicals in regulatory toxicity testing--an integrated approach without additional animal use." Regul Toxicol Pharmacol **63**(2): 321-332.

Sellers, K., N. Deleebeeck, M. Messiaen, M. Jackson, E. Bleeker, D. Sijm and F. van Broekhuizen (2015). Grouping Nanomaterials - A

strategy towards grouping and read-across. R. Bilthoven, the Netherlands, National Institute for Public Health and the Environment (RIVM) & Arcadis**:** 227.

Suter, G., T. Vermeire, W. Munns and J. Sekizawa (2001). Approaches to integrated risk assessment. Chapter II: Framework for the integration of health and ecological risk assessment, WHO.

Suter, G. W., T. Vermeire, W. R. Munns and J. Sekizawa (2003). "Framework for the Integration of Health and Ecological Risk Assessment." Human and Ecological Risk Assessment: An International Journal **9**(1): 281-301.

Teeguarden, J. G., Y.-M. Tan, S. W. Edwards, J. A. Leonard, K. A. Anderson, R. A. Corley, M. L. Kile, S. M. Simonich, D. Stone, R. L. Tanguay, K. M. Waters, S. L. Harper and D. E. Williams (2016). "Completing the Link between Exposure Science and Toxicology for Improved Environmental Health Decision Making: The Aggregate Exposure Pathway Framework." Environmental Science & Technology.

Tollefsen, K. E., S. Scholz, M. T. Cronin, S. W. Edwards, J. de Knecht, K. Crofton, N. Garcia-Reyero, T. Hartung, A. Worth and G. Patlewicz (2014). "Applying Adverse Outcome Pathways (AOPs) to support Integrated Approaches to Testing and Assessment (IATA)." Regul Toxicol Pharmacol **70**(3): 629-640.

Tulve, N., M. Olsen, M. Broder, M. Firestone, E. Mulford, L. Sheldon, C. Stahl and V. Zartarian (2016). Guidelines for Human Exposure Assessment (*peer review draft*), Risk Assessment Forum U.S. Environmental Protection Agency**:** 213.

US EPA (1996). Fate, Transport and Transformation Test Guidelines Oppts 835.1220 Sediment and Soil Adsorption/Desorption Isotherm. USA, US EPA**:** 20.

US EPA (2012). Benchmark Dose Technical Guidance. Washington DC,, US EPA**:** 99.

US EPA (2012). EPI Suite™. Washington, DC, U.S.A., U.S. Environmental Protection Agency (EPA) Office of Pollution Prevention Toxics and Syracuse Research Company (SRC).

van de Meent, D. and J. H. M. de Bruijn (2007). Environmental exposure assessment. Risk Assessment of Chemicals: An Introduction. 2nd edition. C. J. van Leeuwen and T. Vermeire. Dordrecht, Springer**:** 159-193.

van der Ven, L. T. M., W. G. Schoonen, R. M. Groot, F. den Ouden, H. J. Heusinkveld, E. P. Zwart, H. M. Hodemaekers, E. Rorije and J. de Knecht (2020). "The effects of aliphatic alcohols and related acid metabolites in zebrafish embryos - correlations with rat developmental toxicity and with effects in advanced life stages in fish." Toxicol Appl Pharmacol **407**: 115249.

Verhaar, H. J. M., W. de Wolf, S. Dyer, K. C. H. M. Legierse, W. Seinen and J. L. M. Hermens (1999). "An LC50 vs Time Model for the Aquatic Toxicity of Reactive and Receptor-Mediated Compounds. Consequences for Bioconcentration Kinetics and Risk Assessment." Environmental Science & Technology **33**(5): 758-763.

Vermeire, T., R. McPhail and M. Waters (2001). D. Organophosphorous pesticides in the environment. World Health Organization. IV. Meeting Report of the International Workshop on Approaches to Integrated Risk Assessment.

Vermeire, T., W. R. Munns, J. Sekizawa, G. Suter and G. Van der Kraak (2007). "An Assessment of Integrated Risk Assessment." Human and Ecological Risk Assessment: An International Journal **13**(2): 339-354.

Voutchkova, A. M., J. Kostal, K. A. Connors, B. W. Brooks, P. T. Anastas and J. B. Zimmerman (2012). "Towards rational molecular design for reduced chronic aquatic toxicity " Green Chemistry **14**: 1001-1008.

Voutchkova, A. M., J. Kostal, J. B. Steinfeld, J. W. Emerson, B. W. BBrooks, P. T. Anastas and J. B. Zimmerman (2011). "Towards rational molecular design: derivation of property guidelines for reduced acute aquatic toxicity " Green Chemistry **13**: 2373-2379.

Voutchkova, A. M., T. G. Osimitz and P. T. Anastas (2010). "Toward a Comprehensive Molecular Design Framework for Reduced Hazard." Chemical Reviews **110**(10): 5845-5882.

Wilks, M. F., N. Roth, L. Aicher, M. Faust, P. Papadaki, A. Marchis, M. Calliera, A. Ginebreda, S. Andres, R. Kühne and G. Schüürmann (2015). "White paper on the promotion of an integrated risk assessment concept in European regulatory frameworks for chemicals." Science of The Total Environment **521–522**: 211-218.

Zweers, P. G. P. C. and T. G. Vermeire (2007). Environmental exposure assessment. Risk Assessment of Chemicals: An Introduction. 2nd edition. C. J. van Leeuwen and T. Vermeire. Dordrecht, Springer**:** 357-374.

Annex 1 Endpoints in human health hazard assessment

Type of endpoint	**Nano-materials**	**Industrial chemicals**	**Biocides**	**Plant Protection Products**	**Human medicinal products**	**Food and feed**	**Cosmetics**
Irritation – cutaneous – *in vitro*	Under focus	≥1 t/y				Partly covered by the industrial chemicals' (REACH) for production Not defined yet as for their use	
Irritation – cutaneous – *in vivo*		≥10 t/y					
Irritation – ocular – *in vitro*		≥1 t/y					
Irritation – ocular – *in vivo*		≥10 t/y					
Corrosion	Under focus	≥1 t/y					
Sensitization – cutaneous	Under focus	≥1 t/y					
Mutagenicity (*in vitro*, bacterial genic mutations)	Under focus	≥1 t/y					
Mutagenicity (*in vitro*, mammal or micronuclei cell cytogenicity)	Under focus	≥10 t/y					
Mutagenicity (*in vitro*, mammal cell genic mutations)	Under focus	≥10 t/y					
Toxicokinetic	Under focus	≥10 t/y					When relevant
Acute oral toxicity	Under focus	≥1 t/y					
Acute inhalation toxicity	Under focus	≥10 t/y					
Acute cutaneous toxicity	Under focus	≥10 t/y					
Repeated dose toxicity (28 days, one species)	Under focus	≥10 t/y					

Type of endpoint	**Nano-materials**	**Industrial chemicals**	**Biocides**	**Plant Protection Products**	**Human medicinal products**	**Food and feed**	**Cosmetics**
Repeated dose toxicity (90 days, one species, rodent)	Under focus	≥100 t/y					
Reproductive toxicity (screening, one species)		≥10 t/y					
Reproductive toxicity (prenatal development, one species)		≥100 t/y					
Reproductive toxicity (2 generations, one species)		≥100 t/y					
Reproductive toxicity (development, one species)		≥1000 t/y					
Carcinogenicity	Under focus	≥1000 t/y					When relevant
Neurotoxicity							
Immunotoxicity							
Dermal absorption							
In vitro digestion							
ADME (TK + metabolism)		Under focus					
Endocrine data		Under focus	Under focus	Under focus	Under focus		

Red field = endpoints deemed mandatory in regard to European legislation;
"under focus" refers to the endpoints that are not identified as mandatory in regard to European regulation but which are most likely to be included in the coming years.

Annex 2 Overview of the in vitro toxicity test used for human health risk assessment, including the toxicological endpoints measured and the test species used

Type of study / route / *in vivo* or *in vitro* / OECD TG	**Toxicological endpoint**	**Species**
In vitro skin corrosion: The Transcutaneous Electrical Resistance (TER using rat skin) test OECD TG 430	skin corrosion: necrosis, ulcers, bleeding; discoloration, scars, alopecia, etc.	rat skin
In vitro skin corrosion: Human Skin Model tests OECD TG 431	skin irritation or corrosion: decrease in cell viability, inflammatory mediators release	human skin
In vitro Membrane Barrier Test Method for Skin Corrosion OECD TG 435	skin irritation or corrosion	
In vitro Skin Irritation Reconstructed Human Epidermis Test Method OECD TG 439	skin irritation, inflammatory mediators release, cell viability	human skin
Bovine Corneal Opacity and Permeability Test Method for Identifying Ocular Corrosives and Severe Irritants OECD TG 437	irritation or corrosion to cornea and conjunctiva	bovine eye
Isolated Chicken Eye Test Method for Identifying Ocular Corrosives and Severe Irritants OECD TG 438	irritation or corrosion to cornea and conjunctiva	Chicken Eye
Dermal sensitisation LLNA Local Lymph Node Assay OECD TG 429, 442A, 442B	primary proliferation of lymphocytes; clinical signs, either of local irritation (ear erythema or ear thickness) at the application site or of systemic toxicity	mouse (CBA/Ca or CBA/J strain)
Genotoxicity (mutagenicity) Ames Test, Bacterial reverse	mutagenicity	Bacterial cells *Salmoella Typhimurium* spp)

Type of study / route / *in vivo* or *in vitro* / OECD TG	**Toxicological endpoint**	**Species**
mutation test OECD TG 471 (*in vitro*)		
Genotoxicity (mutagenicity) *In vitro* mammalian chromosome aberration test OECD TG 473	structural and numerical chromosome aberrations, clastogenicity, aneugenicity	established cell lines, cell strains or primary cell cultures or lymphocytes
Genotoxicity (clastogenicity and aneugenicity) *In vitro* micronucleus test OECD TG 487	Structural and numerical chromosome (or chromatid) aberrations	Cell cultures of human or mammalian origin
Genotoxicity, (mutagenicity) *In vitro* mammalian cell gene mutation test – hprt test OECD TG 476	gene mutations in the hprt gene of established cell lines	mammalian cell lines (L5178Y mouse lymphoma cells; CHO; AS52 and V79 lines of Chinese hamster cells; TK6 human lymphoblastoid cells
Genotoxicity Mouse spot test OECD TG 484	detection of presumed somatic mutations in foetal cells	appropriate strains of mated mice
Endocrine disruption H295R Steroidogenesis Assay OECD TG 456	Endocrine disruption; *in vitro* screen for chemical effects on steroidogenesis, specifically the production of 17ß-estradiol and testosterone	human H295R adreno-carcinoma cell line
mechanistic study The Stably Transfected Human Estrogen Receptor-alpha Transcriptional Activation Assay for Detection of Estrogenic Agonist-Activity of Chemicals OECD TG 455		hERα-HeLa-9903 cell line

Annex 3 Endpoints in environmental risk assessment

Compartment	Type of endpoint	Industrial chemicals	Biocides*	Plant Protection Products*	Human medicinal products	Veterinary medicinal products	Food and feed	Cosmetics
Water (default freshwater, marine and/or brackish water if relevant)	Acute toxicity – initial test		AS core	AS core				
	Acute toxicity – algae/aquatic plants	≥1 t/y	AS core	AS core		Phase IIA		
	Acute toxicity – invertebrates	≥1 t/y	AS core	AS core		Phase IIA		
	Acute toxicity – fish	≥10 t/y	AS core	AS core		Phase IIA		
	Toxicity – microorganisms	≥10 t/y	AS core		Phase IIA			
	Chronic toxicity – algae	≥1 t/y	AS additional	AS core	Phase IIA	Phase IIB		
	Chronic toxicity – invertebrates	≥100 t/y	AS additional	AS core	Phase IIA	Phase IIB		
	Chronic toxicity – fish	≥100 t/y	AS additional	AS core	Phase IIA	Phase IIB		
	Bioaccumulation – aquatic organisms	≥100 t/y	AS core	AS core	Phase IIB	Phase IIB	Not yet defined	Description of all available ecological and environmental effects of substance/ compound/ mixture
	Higher Tier – micro-/mesocosm, field studies			if triggered				

Compartment	Type of endpoint	Industrial chemicals	Biocides*	Plant Protection Products*	Human medicinal products	Veterinary medicinal products	Food and feed	Cosmetics
Sediment (default freshwater, marine and/or brackish water if relevant)	Chronic toxicity – invertebrates	≥1000 t/y	AS additional	AS core (if substance partitions to sediment)	Phase IIB	Phase IIB		
Soil	Acute toxicity – plants	≥ 100 t/y	AS additional	AS core		Phase IIA		
	Acute toxicity – non target invertebrates	≥ 100 t/y	AS additional	AS core	Phase IIB			
	Toxicity – non target microorganisms	≥100 t/y	AS additional	AS core		Phase IIA		
	Chronic toxicity – plants	≥1000 t/y	AS additional	AS core	Phase IIB	Phase IIB		
	Chronic toxicity – non target invertebrates (e.g. Collembola)	≥1000 t/y	AS additional	AS core	Phase IIB	Phase IIA		
	Higher Tier Plants (semi-field, field)			if triggered				
	Bioaccumulation - invertebrates	if triggered	AS additional			Phase IIB		
Other	Acute toxicity – honeybees or other arthropods		AS additional	AS core				
	Dietary - honeybees			AS core				

Compartment	Type of endpoint	Industrial chemicals	Biocides*	Plant Protection Products*	Human medicinal products	Veterinary medicinal products	Food and feed	Cosmetics
	Effects – non-target arthropods		AS additional	AS core		Phase IIA		
	Higher tier – honey bees or other arthropods (e.g. aged-residues, semi-field, field)			if triggered				
	Avian - Acute oral		AS additional	AS core				
	Avian - Short-term toxicity		AS additional	AS core				
	Avian - Chronic toxicity or reproductive toxicity	≥1000 t/y	AS additional	AS core				
	Mammals – Acute oral	≥1 t/y	AS core	AS core				
	Endocrine data (2 amphibian & 3 fish tests)	See ED case study						

Red field = endpoints deemed mandatory in regard to European legislation;
Orange field = endpoints deemed mandatory if triggered in regard to European legislation;
Yellow field = endpoints deemed optional/indicative in regard to European legislation;
AS core = Active substance core data;
AS additional = Active substance additional data that is biocidal product specific;
* = biocides and pesticides are also tested as products (in addition to active substances) to account for formulation effects (e.g. synergy), this is not included in the above table.

Annex 4 A pathway based integration of human health and environmental testing and assessment of endocrine disrupting chemicals

Risk assessment of endocrine disrupting chemicals

A pathway based integration of human health and environmental testing and assessment is an important issue for identification and risk assessment of endocrine disrupting chemicals (EDCs). Although biological pathways, e.g. estrogen receptor (ER), are relatively conservative among vertebrates, there are great challenges in terms of interpretation of chemical effects across species. This is partly due to differences in ADME (adsorption, distribution, metabolism and elimination) of chemicals across species, e.g. water exposure of fish versus oral exposure of rodents. Another reason is that many pathways regulate the same endpoint, e.g. pathways of ERs and peroxisome proliferator-activated receptors (PPARs) regulate reproduction. Differences in basic biology, e.g. live-bearing versus egg-laying vertebrates, also lead to challenges for a pathway based interpretation of chemical effects across species. To integrate information of human health and environmental tests across species, it is essential to identify and evaluate endpoints that have been directly linked to certain pathways. As a part of the project of IRAC (Integration and innovation of ecological and human health Risk Assessment: Connecting concepts and cases), an inventory is made for the EATS (estrogen, androgen, thyroid, and steroidogenesis) pathways because only these pathways can be captured in the current OECD test guidelines.

Inventory of pathway-related tests across species
This annex focuses on the regulatory tests. For EDCs, these tests have been summarized in the OECD Conceptual Framework (CF) for Testing and Assessment of Endocrine Disrupters.

This CF lists the OECD Test Guidelines and standardized test methods available, under development or proposed that can be used to evaluate chemicals for endocrine disruption. These tests are organized into five levels (Table 6).

Level 1 includes existing data and non-test information. All existing information that is not included at levels 2 to 5 should be collated at this level. Such information include the physical-chemical properties of a chemical of interest; *in silico* predications like QSARs, read across; and ADME predications; Data from literature studies on (eco)toxicological tests that are not included at levels 2-5 can be included at Level 1. Information at level 1 can be used for a weight of evidence analysis. This paper will not further discuss this level for the integration of human health and environmental tests because the specific assays and pathways are listed in levels 2 to 5.

Level 2 contains *in vitro* assays providing data about selected endocrine mechanisms or pathways. As shown in Table 6, test guidelines are available for estrogen receptor (ER) binding and transactivation assays

as well as for the steroidogenesis (S) *in vitro* assay. Test guidelines for androgen receptor (AR) binding and transactivation assays are under development and the test guideline for thyroid receptor transactivation is not available at this stage. Another *in vitro* test the MCF-7 cell proliferation assay has been listed at level 2 of the CF. The development of test method, however, has been stopped from the OECD program. Overall, the current *in vitro* tests available are only for elucidating EAS pathways. The other pathways may be considered when the test methods are adopted by the OECD.

It is also noted that all tests use only mammalian cell lines and receptors. No other species e.g. fish, are available. This raises a question whether information derived from mammalian systems can be equally used for non-mammalian vertebrate and invertebrate systems. The same is true for extrapolation of non-mammalian information to mammalians when integration of information of human health and environmental tests is needed.

Level 3 includes *in vivo* assays providing data about selected endocrine mechanisms or pathways. Two assays, the Uterotrophic assay and the Hershberger assay, are available for mammalians, with the former for detecting the ER-mediated pathway and the latter for the AR-mediated pathway. Both assays are considered for indicating MOAs but not for risk assessment because the experimental animals need to be operated in these assays. Five tests are available for non-mammalians, including 2 amphibian tests and 3 fish tests. Two amphibian tests are used for detecting chemicals interfering with the Thyroid hormone pathway (T). The test, xenopus embryo thyroid signaling assay, is currently in Phase II validation. Three fish assays have been adopted by the OECD, with two assays of TG229 and TG230 for indicating EAS pathways; one assay of GD148 for the anti-androgenic pathway. The latter fish assay was developed due to the fact that the first two fish assays cannot consistently detect fish biomarker changes induced anti-androgenic chemicals. Different from the mammalian tests at Level 3, the adopted non-mammalian test methods are performed in intact amphibians and fish. There are some population-relevant apical endpoints included in the test methods. Whether or not the results of such tests can be used for risk assessment and classification depends on the test design, e.g. the number of animals and test concentrations. The current test design of the OECD protocols is purely for the screening purpose. In summary, two mammalian tests are available for detecting E and A pathways, respectively; 5 non-mammalian tests are listed, with 2 amphibian tests for elucidating the T pathway, and 3 fish tests for EAS pathways. Results of some non-mammalian assays may be used for risk assessment and classification, if the test design is adapted.

In vivo assays providing data on adverse effects on endocrine relevant endpoints are listed at Level 4. There are 10 tests included for mammalians and 12 tests for non-mammalians. Except a few non-mammalian tests, all assays at this level include apical endpoints and are initially designed for hazard and risk assessment. Due to the great impact of EDCs, OECD has been developing tests or updating existing test methods for detecting endocrine activities. For mammalian tests, the 28-day Repeated Dose Toxicity Test (OECD TG 407) and the

Reproduction/Developmental Toxicity Screening Test (TG421) have been updated in 2008 and 2015, respectively. This update includes parameters suitable to detect EATS mediated activity. However, the sensitivity of updated assays may not be sufficient to identify all EATS-mediated EDCs (OECD 2012). The update of TG414 by including endocrine sensitive parameters is currently under development. Three assays do not have OECD test guidelines but have been included in the US endocrine disruptor screening program (EDSP). These three assays are male pubertal assay, female purertal assay and intact adult male endocrine screening assay, which can detect potential of chemical interacting with EATS pathways (OECD 2012). Other tests at level 4 have limited capacity to detect endocrine modes of action (MOAs) including TG409, TG451-3, and TG 426. Different from mammalian tests, fish and amphibian tests (TG234 and TG241) are initially designed to detect both adverse effects and endocrine MOAs. Therefore, both apical and biomarker endpoints have been validated during the test method development. TG234 can detect chemicals interfering with EAS pathways; TG241 can detect the T pathway interference as well as EA pathways. Different from fish and amphibians, TG206 is only used for detecting adverse effects. There are two mollusc assays that are just adopted this year. These two assays, however, include only apical endpoints due to lack of knowledge on invertebrate endocrinology. Similarly, the other assays on invertebrates, except the daphnia test, included at level 4 can only detect adverse effects. The Daphnia test may detect chemicals interfering with the Juvenile hormone pathway (Dang et al., 2012). However, this test for detecting this pathway has not yet been validated. Taken together, some tests at Level 4 can detect both adverse effects and the interference of EATS pathways. Other tests, however, can only detect adverse effects of a chemical of interest.

Level 5 consists of *in vivo* assays providing more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organisms. There are two tests for mammalians, the Extended One-Generation Reproductive Toxicity Study (EOGRTS) and the two-generation toxicity test (Table 6). The two-generation toxicity test (TG416) was initially designed to detect adverse effects of a chemical and was not considered as a sensitive test method for detecting endocrine activity. The most recent version of this test was adopted in 2001. In this version, some endocrine endpoints like estrous cyclicity and primordial follicle counts were included. However, the most recent version of TG 416 does not include some endocrine sensitive endpoints such as nipple retention. No further update is planned in the OECD. Instead, OECD has developed an EOGTRS, in which endocrine sensitive endpoints in the juvenile and adult F1 have been included. The EOGRTS is preferable for detecting endocrine disruption. The basic study design of the EOGRTS focuses on evaluation of the fertility of parental animals and of defined parameters on postnatal development of F1 animals until adulthood. It does not include mating of F1 animals (producing F2 generation) or cohorts for DNT (developmental neurotoxicity) or DIT (developmental immunotoxicity). Conditions for triggering an extension of F2, DNT and DIT may vary with the different legal frameworks. Both TG416 and TG443 are considered to detect the interference of EATS pathways. For non-mammalian tests, many tests

have been included in the CF. However, a majority of them have not yet been developed. Medaka multi-generation test has been changed into medaka extended one generation test (MEOGRTS), which has been adopted in 2015. In addition to adverse effects, MEORGTS can detect changes in EAS pathways by including fish biomarkers. Mysid and Copepod assays have been removed from the OECD development program because the leading country stopped the project. TG233 for chironomid life cycle test is included at level 5. However, it is unknown whether this assay can detect relevant endocrine MOAs because of the lack of knowledge in invertebrate endocrinology. Again, in addition to adverse effects, tests included at level 5 can detect chemical-induced changes in EATS pathways.

Taken together, tests listed at levels 2 to 5 of the OECD CF are considered as the most important test methods for detecting endocrine disruption (OECD 2012). The OECD CF includes standard test methods that can be used for detecting both adverse effects, endocrine MOAs and the causal relationship between adversity and MOAs. At this stage, however, these tests cover only EATS pathways. Test methods for other pathways that are relevant to EDC-induced diseases, e.g. obesity and diabetes, are not included in the CF because such a type of test methods has not yet been validated. This CF is a toolbox but not a testing strategy. This means that any test at any level can be conducted and it is not necessary to follow the CF in a linear manner.

Table 6 OECD Conceptual Framework for Testing and Assessment of Endocrine Disruptors

Levels	Tests	TG	MOAs
Level 1 Existing data and non-test information	• Physical & chemical properties, e.g., MW, reactivity, volatility, biodegradability. • All available toxicological data from standardized or non-standardized tests. • Read across, chemical categories, QSARs and other *in silico* predictions, and ADME model predictions.		
Level 2 *In vitro* assays providing data about selected endocrine mechanism(s)/pathways(s)	• Estrogen receptor binding affinity • Estrogen receptor transactivation • Androgen receptor binding affinity • Androgen transactivation • thyroid transactivation • Steroidogenesis *in vitro* • MCF-7 cell proliferation assays • Other assays as appropriate	493 455 dev. dev. n.a. 456 stop	E E A A T S E
Level 3 *In vivo* assays providing data about selected endocrine mechanism(s)/ pathway(s)	**Mammalian Toxicology** • Uterotrophic assay • Hershberger assay **Non-Mammalian Toxicology** • Xenopus embryo thyroid signalling assay • Amphibian metamorphosis assay • Fish Reproductive Screening Assay	 440 441 Dev 231 229 230	 E A T T EAS EAS

Levels	Tests	TG	MOAs
	• Fish Screening Assay • Androgenized female stickleback screen	GD148	A
Level 4 *In vivo* assays providing data on adverse effects on endocrine relevant endpoints	**Mammalian Toxicology** • Repeated dose 28-day study • Repeated dose 90-day study • 1-generation reproduction toxicity study • Male pubertal assay • Female pubertal assay • Intact adult male endocrine screening assay • Prenatal developmental toxicity study • Chronic toxicity and carcinogenicity studies • Combined 28-day/reproductive screening assay • Developmental neurotoxicity **Non-Mammalian Toxicology** • Fish sexual development test • Fish Reproduction Partial Lifecycle Test • Larval Amphibian Growth & Development Assay • Avian Reproduction Assay • Mollusc Partial Lifecycle Assays • Chironomid Toxicity Test • *Daphnia* Reproduction Test (with male induction) • Earthworm Reproduction Test • Enchytraeid Reproduction Test • Sediment Water *Lumbriculus* Toxicity Test Using Spiked Sediment • Predatory mite reproduction test in soil • Collembolan Reproduction Test in Soil	 407 408 415 EPA EPA EPA 414 451-3 421-2 426 234 n.a. 241 206 ado 218-9 211 222 220 225 226 232	 EATS - - EATS EATS EATS - - - - EAS - EAT - - - - - - - - -
Level 5 *In vivo* assays providing more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism	**Mammalian Toxicology** • Extended one-generation reproductive toxicity study • 2-Generation reproduction toxicity study **Non-Mammalian Toxicology** • Fish Life Cycle Toxicity Test • Medaka Extended one-generation Test • Avian 2 generation reproductive toxicity assay • Mysid Life Cycle Toxicity Test • Copepod Reproduction and Development Test • Sediment Water Chironomid Life Cycle Toxicity Test	 443 416 n.a. 240 n.a. stop stop 233	 EATS EATS - EAS - - - -

Inventory of pathway-related endpoints across species in vivo
The OECD guidance document (GD, 150) on standardized test guidelines for evaluating chemicals for endocrine disruption summarized possible endpoints and their applicability for identifying endocrine disrupting mechanisms and/or effects resulting from EATS pathways. These endpoints were derived from validation studies of TG407. Based on the direction of changes in relevant endpoints, the possible MOAs could be indicated. Extrapolations have been made across similar endpoints in different studies. For example, TG 416 has not been validated for thyroid-related activities but it is assumed that thyroid changes in OECD TG 416 would be similar to those of TG407. As the same endpoint may be modulated via different MOAs, the indication of endocrine MOAs has some uncertainties in these in vivo mammalian tests. Endpoints for the majority of in vivo tests from level 3 to 5 have not yet received full validation for endocrine MOAs (Table 7).

In contrast to mammalian tests, all endpoints which are indicative of endocrine MOAs have been validated in fish and amphibians. Overall, limited number of biomarkers are used for indication of endocrine MOAs (Table 8). Vitellogenin (VTG), an egg yolk precursor protein, is normally produced by the liver of female oviparous animals in response to circulating endogenous estrogens. It is also detected in heart, spleen, kidney, skin, muscle, gill, eye and brain tissues in fish (Dang, 2016). A low level of VTG can be detected in the plasma of male and immature fish because of low circulating estrogen stimulation. In the presence of estrogens or EDCs, however, the liver is induced to synthesize and secrete VTG (OECD, 2012). Early studies focused on the effects of EDCs on VTG protein levels in adult fish. Recent studies showed that changes in VTG protein and mRNA levels could be observed in different stages of fish exposed to EDCs. In addition to fish, this endpoint is considered indicative of endocrine |MOAs in amphibians. This endpoint is included in test guidelines 229, 230, 234, 240 and 241. According to the OECD test guidelines, changes in VTG can be used for the detection of chemicals interfering with EAS pathways (OECD, 2012). Secondary sex characteristics (SSC) in male fathead minnow and medaka are externally visible, quantifiable and responsive to circulating levels of androgens or EDCs. Stimulated by androgens or EDCs, females could develop male SSC. The main indicators of androgenic exposure are the number of nuptial tubercles located on the snout of the female fathead minnow and the number of papillary processes in female medaka. This endpoint has been included in test guidelines 229, 230, 234 and 241. Similar to the endpoint VTG, changes in SSC are used to indicate chemicals interfering with EAS pathways in both fathead minnow and medaka (OECD, 2012). The phenotypic sex ratio and genetic sex ratio are two types of sex ratio in fish often referred to in the literature. The phenotypic sex ratio is determined in individual fish via the histological examination of the gonads and is defined as female, male, intersex (both oocytes and spermatogenetic cells in one gonad) or undifferentiated. The phenotypic sex ratio can be determined in all three small model fish species (OECD, 2012). In contrast, the genetic sex marker is available in medaka but not in zebrafish and fathead minnow. The genetic sex ratio of medaka can be examined via genetic markers, i.e. the female XX or male XY genes, or the Y-linked DM domain gene (DMY). Different from VTG and SSC, sex ratio is considered as a

biomarker and a population relevant endpoint that can indicate MOAs and can be used for risk assessment and for classification of chemicals (OECD, 2012). This endpoint is considered indicative of endocrine MOAs both in fish and amphibians. Sex ratio has been included in test guidelines 234, 240, and 241. Similar to VTG and SSC, changes in sex ratio are used to indicate chemicals interfering with EAS pathways (OECD, 2012). In addition to biomarkers, changes in histology of target organs e.g. thyroid and gonad, is often used as indication of endocrine MOAs in amphibians (Table 8).

Integration of human health and environmental tests for endocrine MOAs

As mentioned above, both mammalian and non-mammalian tests can capture EATS pathways by using different endpoints. Integration of results derived from mammalian and non-mammalian tests is necessary because of the broad similarity of endocrine systems across the vertebrates. Table 9 summarizes tests that could be a good foundation for the same pathways. Integration of these mammalian and non-mammalian tests provides a good basis for a weight of evidence analysis of EATS pathways.

Conclusion

Currently available standardised guideline tests are adequate in most cases for EATS pathways in mammalian and non-mammalian vertebrates. Extrapolation of information from mammalian to non-mammalian vertebrates and vice versa is possible. Integration of information derived from these two groups of vertebrates provides a good basis for indication of endocrine MOAs of EATS pathways. General knowledge of the endocrine system in invertebrates is currently limited. It is therefore difficult to integrate information derived from invertebrates with that from vertebrates including mammalians. Standard test methods for other pathways like peroxisome proliferator-activated receptor (PPAR) signalling pathway essential for regulating obesity have not yet been developed. Integration of mammalian and non-mammalian information for these pathways are not possible at this stage.

s

Table 7 Endpoints in in vivo mammalian assays relevant to EATS pathways (based on OECD GD 150)

	OECD CF	OECD TG	endpoints	pathways	TR-pathway	Steroidogenesis pathway
level 3	Uterotrophic assay	440	uterine weight	E and aromatisable androgens		
	Hershberger assay	441		ventral prostate, seminal vesicles, LABC, cowpers glands, glans penis	Optional: changes in serum T3 and T4, histopathologica changes in thyroid	
level 4	repeated dose toxicity study (28 days)	407	histopathologic and weight changes in ovary, uterus/cervix, vagina; testes, epididymides, prostate + seminal vesicles with coagulating glands. Optional endpoints: weight of uterus and ovaries; changes in vaginal smears; histopathologic changes in mammary glands (males).	histopathologic changes in ovary, uterus/cervix, vagina; weight of testes, prostate + seminal vesicles with coagulating glands. histopathologic changes in testes, epididymides, Optional endpoints: ovary weight, changes in vaginal smears. histopathologic changes in mammary glands.	Liver weight in combination with other thyroid-related endpoints, histopathologic changes in Thyroid (follicular cell, height increase & colloid area decrease) Optional: Serum T3 and T4, TSH, thyroid weight.	Histopathologic changes in ovary, uterus/cervix, vagina. weight of prostate + seminal vesicles with coagulating glands. Optional: Uterine and ovary weight, changes in vaginal smears. Histopathologic changes in mammary gland.

	OECD CF	OECD TG	endpoints	pathways	TR-pathway	Steroidogenesis pathway
	male pubertal assay	EPA	Assay is not designed to detect this modality but the following changes may occur: age at preputial separation; weight of seminal vesicles (+ coagulating glands); ventral prostate, dorsolateral prostate, LABC, epididymides; testis weight; histopathologic changes in testes epididymides; serum testosterone	age at preputial separation; weight of seminal vesicles, ventral prostate, dorsolateral prostate, LABC, epididymides, testis weight, histopathologic changes in testes, epididymides, serum testosterone	thyroid weight, liver weight, histopathologic changes in thyroid, serum T4, TSH	preputial separation, weight of seminal vesicles, ventral prostate, dorsolateral prostate, LABC, epididymides, histopathologic changes in testes epididymides, serum testosterone
	female pubertal assay	EPA	weight of uterus & ovaries; histopathologic changes in uterus & ovaries; age at first estrus; changes in estrus cyclicity.	Assay is not designed to detect this modality but the following changes may occur: age at vaginal opening, weight of uterus and ovaries, histopathologic changes in uterus and ovaries, age at first estrus, estrus cyclicity	thyroid weight, liver weight, histopathologic changes in thyroid, serum T4, TSH	age at vaginal opening, weight of uterus and ovaries, histopathologic changes in uterus and ovaries, estrus cyclicity

	OECD CF	OECD TG	endpoints	pathways	TR-pathway	Steroidogenesis pathway
level 5	Extended One-Generation Reproductive Toxicity Study (basic)	443	Change in AGD, estrus cyclicity, age at vaginal opening, age at preputial separation, weights of uterus, ovaries, testes, epididymides, prostate, seminal vesicles (+ coagulating glands); histopathologic changes in vagina, uterus (+ cervix), ovaries, testis, epididymis, prostate, seminal vesicles and coagulating glands. Changes in sperm parameters; Genital abnormalities; histopathologic changes (proliferative) in mammary glands.	AGD, age at preputial separation, genital abnormalities; weights of uterus, ovaries, testes, epididymides, prostate, seminal vesicles (+coagulating glands); histopathologic changes in the above organs and in mammary glands, changes in sperm parameters, nipple retention.	Liver weight in combination with other thyroid-related endpoints, histopathologic changes in Thyroid (follicular cell, height increase & colloid area decrease) Optional: Serum T3 and T4, TSH, thyroid weight.	Change in AGD, estrus cyclicity, age at vaginal opening, age at preputial separation, weights of uterus, ovaries, testes, epididymides, prostate, seminal vesicles (+ coagulating glands); histopathologic changes in vagina, uterus (+ cervix), ovaries, testis, epididymis, prostate, seminal vesicles and coagulating glands. Changes in sperm parameters; Genital abnormalities; histopathologic changes (proliferative) in mammary glands.

	OECD CF	OECD TG	endpoints	pathways	TR-pathway	Steroidogenesis pathway
	two-generation toxicity test	416	Change in AGD, estrus cyclicity, age at vaginal opening, age at preputial separation, weights of uterus, ovaries, testes, epididymides, prostate, seminal vesicles (+ coagulating glands); histopathologic changes in vagina, uterus (+ cervix), ovaries, testis, epididymis, prostate, seminal vesicles and coagulating glands. Changes in sperm parameters.	Change in AGD, estrus cyclicity, age at vaginal opening, age at preputial separation, weights of uterus, ovaries, testes, epididymides, prostate, seminal vesicles (+ coagulating glands); histopathologic changes in vagina, uterus (+ cervix), ovaries, testis, epididymis, prostate, seminal vesicles and coagulating glands. Changes in sperm parameters.	thyroid weight; liver weight in combination with other thyroid-related endpoints; histopathologic changes in thyroid (follicular cell height increase & colloid area decrease)	Change in AGD, estrus cyclicity, age at vaginal opening, age at preputial separation, weights of uterus, ovaries, testes, epididymides, prostate, seminal vesicles (+ coagulating glands); histopathologic changes in vagina, uterus (+ cervix), ovaries, testis, epididymis, prostate, seminal vesicles and coagulating glands. Changes in sperm parameters.

Table 8 Endpoints in in vivo non-mammalian assays relevant to EATS pathways

OECD CF	test	OECD TG	Endpoints	E	A	S	T
level 3	fish reproduction screening assay	229	VTG, SSC	x	x	x	
	fish screening assay	230	VTG, SSC	x	x	x	
	Androgenized female stickleback screen	GD 148	Spiggin		x		
	Amphibian metamorphosis assay	231	Thyroid Gland Histology, developmental stage and body length/weight				x
Level 4	Fish sexual development test	234	VTG, sex ratio	x	x	x	
	Larval Amphibian Growth & Development Assay	241	thyroid histopathology, gonad and gonad duct histopathology, abnormal development, plasma vitellogenin (optional), and genotypic/phenotypic sex ratios	x	x		x
level 5	Medaka extend one generation test	240	VTG, SSC, sex ratio	x	x	x	

Table 9 Integration of mammalian and non-mammalian tests for EATS pathways

Pathway	***In vitro***	***In vivo* mammalian**	***In vivo* non-mammalian**
Estrogen	ER binding, transactivation	male or female pubertal assay, OECD TG 407, 421, 416, 440 and 443	OECD TG 229, 230, 234, 240 and 241
Androgen	AR binding, transactivation	male or female pubertal assay, OECD TG 407, 421, 416, 440 and 443	OECD TG 229, 230, 234, 240 and 241
Steroidogenesis	Steroidogenesis assay	male or female pubertal assay, OECD TG 407, 421, 416 and 443	OECD TG 229, 230, 234, 240 and 241
thyroid		male or female pubertal assay, OECD TG 407, 421, 416 and 443	OECD TG 230 and 241

RIVM
Committed to *health and sustainability*