European Food Safety Authority

EFSA Journal 2013;11(3):3132

SCIENTIFIC OPINION

Scientific Opinion on the hazard assessment of endocrine disruptors: Scientific criteria for identification of endocrine disruptors and appropriateness of existing test methods for assessing effects mediated by these substances on human health and the environment[1]

EFSA Scientific Committee[2, 3]

European Food Safety Authority (EFSA), Parma, Italy

ABSTRACT

Upon request of the European Commission, the Scientific Committee (SC) of the European Food Safety Authority reviewed existing information related to the testing and assessment of endocrine active substances (EASs) and endocrine disruptors (EDs). This work was conducted by a working group of experts in endocrinology, risk assessment and toxicology, together with observers from other EU agencies, namely EMA, ECHA and EEA. To distinguish between EDs and other groups of substances with different modes of action, it was concluded that an ED is defined by three criteria: the presence of i) an adverse effect in an intact organism or a (sub)population; ii) an endocrine activity; and iii) a plausible causal relationship between the two. As scientific criteria for adversity have not been generally defined, specific criteria for endocrine disrupting effects could not be identified. Hence, expert judgement is required to assess on a case-by-case basis the (eco)toxicological relevance of changes at the molecular to individual and/or (sub)population level following exposure to an EAS. The SC concluded that a reasonably complete suite of standardised assays for testing the effects of EASs is (or will soon be) available for the oestrogenic, androgenic, thyroid and steroidogenic modalities in mammals and fish, with fewer tests for birds and amphibians. Shortcomings in current tests and for other endocrine modalities and species were reviewed. Critical effect, severity, (ir)reversibility and potency aspects are part of the hazard characterisation of EDs. To inform on risk and level of concern for the purpose of risk management decisions, risk assessment (taking into account hazard and exposure data/predictions) makes best use of available information. Levels of concern are not determined exclusively by risk assessment but also by protection goals set by the risk management.

KEY WORDS

Endocrine active substance (EAS), endocrine disruptor (ED), hazard assessment, test methods, (eco)toxicological risk assessment, adverse effects.

---

[1] On request from the European Commission, Question No EFSA-Q-2012-00760, adopted on 28 February 2013.

[2] Scientific Committee members: Jan Alexander, Diane Benford, Qasim Chaudhry, Anthony Hardy, Michael John Jeger, Robert Luttik, Ambroise Martin, Bernadette Ossendorp, Simon More, Alicja Mortensen, Birgit Noerrung, Joe Perry, Iona Pratt, John Sofos, Josef Schlatter, Kristen Sejrsen. Correspondence: scientific.committee@efsa.europa.eu.

[3] Acknowledgement: The SC wishes to thank the members of the working group on Endocrine Active Substances: Jan Alexander, Jacques Auger, Diane Benford, Susy Brescia, Gisela Degen, Anthony Hardy, Karen Ildico Hirsch-Ernst, Peter Hoet, Susanne Hougaard Bennekou, Robert Luttik, Peter Matthiessen, Wim Mennes, Peter Pärt, Daniel Pickford, Thomas Platzek, Josef Schlatter, Emanuela Testai and Theo Vermeire for the preparatory work on this scientific output; the representatives of the European Commission and EU Agencies: Gabriele Schoening, Niklas Andersson, Jean-Marc Vidal, Dorota Jarosinska, Sharon Munn, Daskaleros Panagiotis, Michael Walsh; the hearing experts: Susan Jobling, Trine Husøy and EFSA staff: Bernard Bottex, Miriam Jacobs, Djien Liem, Reinhilde Schoonjans and Sarah Trattnig for the support provided to this scientific opinion.

Suggested citation: EFSA Scientific Committee; Scientific Opinion on the hazard assessment of endocrine disruptors: scientific criteria for identification of endocrine disruptors and appropriateness of existing test methods for assessing effects mediated by these substances on human health and the environment. EFSA Journal 2013;11(3):3132. [84 pp.] doi: 10.2903/j.efsa.2013.3132. Available online: www.efsa.europa.eu/efsajournal

SUMMARY

Following a request from the European Commission (EC), the Scientific Committee (SC) of the European Safety Authority (EFSA) was asked to deliver a scientific opinion on the hazard assessment of endocrine disruptors (EDs). More specifically the SC was asked to advise on i) the scientific criteria to distinguish between EDs and other groups of substances with different modes of action, ii) the criteria to distinguish between physiological modulation and adverse effects on humans and on the ecosystem as a result of exposure to endocrine active substances (EASs), and iii) to review existing test methods and discuss their appropriateness for the identification and characterisation of effects mediated by EASs.

The SC defines an EAS as a substance having the inherent ability to interact or interfere with one or more components of the endocrine system resulting in a biological effect, but need not necessarily cause adverse effects. Therefore, the SC considers endocrine activity as a collection of modes of action, potentially leading to adverse outcomes, rather than a (eco)toxicological hazard in itself. The SC endorsed the WHO/IPCS 2002 definition of an ED implying that there must be reasonable evidence for a biologically plausible causal relationship between the endocrine activity and the induced adverse effect(s) seen in an intact organism or a (sub)population for a substance to be identified as an ED.

Assessment of adversity is not unique to endocrine related effects. Scientific criteria for assessment of adversity have not been generally defined. In general, but not always, transient, inconsistent and minor fluctuations at the biochemical and molecular level may be considered adaptive, i.e. non-adverse. Changes at the cell-, organ-, organism-, or (sub)population-level resulting in pathology or functional impairment *in vivo*, as well as altered timing of development, may be considered adverse. It is therefore difficult to propose ED-specific criteria for adversity and expert judgement in a weight-of-evidence approach is needed to assess substances for possible endocrine disrupting properties.

The Organisation for Economic Co-operation and Development (OECD) revised Conceptual Framework (CF) provides a guide to the data sources, the OECD test guidelines and standardised test methods available, under development or proposed for the evaluation of chemicals for EASs/EDs. Information on endocrine activity can be obtained from existing information, read-across, *in silico* tools, *in vitro* and *in vivo* screening assays (Levels 1, 2 and 3 of the CF), or from other mechanistic investigations. A prerequisite for an EAS to be regarded as an ED is the need to identify the adverse effect. For this, test methods with apical endpoints (Levels 4 and 5 of the CF) can be used together with existing information, read-across and other *in vivo* (eco)toxicity tests that provide information on apical endpoints. Therefore, in principle, no single assay is likely to provide all the information needed to decide whether a substance is an ED because of the need to provide both mechanistic and apical information.

Taken together, but bearing in mind the recommendations made in this opinion, a reasonably complete suite of standardised assays (for testing the effects of EASs) is (or will soon be) available for the oestrogen, androgen, thyroid, or steroidogenesis (EATS) modalities in mammals and fish, with fewer tests available for birds and amphibians. While downstream effects of disruption of some non-EATS pathways/modalities may be detectable in some of the standardised apical vertebrate assays, it is important to recognise that standardised mechanistic assays for non-EATS modalities relevant to mammals, fish and other vertebrates are not or not yet available. For invertebrates, standardised mechanistic assays are lacking from the OECD testing suite, mainly due to poor understanding of invertebrate endocrinology. Finally, a range of major taxa, e.g. reptiles or echinoderms have not yet been considered by OECD for any endocrine assay development. It is unknown at present whether it will be possible to read-across to untested groups from tests with other taxa.

The SC identified the need for further development of screens and test methods, particularly with regard to non-EATS modalities that may be associated with adverse effects in humans or the environment.

The SC discussed a number of general aspects related to the testing of substances (independently of whether these are EASs or substances with other modes of action):

- Exposure of organs and tissues to certain substances at critical point(s) during their development can result in irreversible change of the organ/tissue. The SC noted that, although some of the tests in the OECD CF do cover exposure during critical periods of development *in utero*, current mammalian tests may not cover effects that might be induced by exposure during fetal or pubertal development, but may emerge during later life stages. Fish lifecycle tests cover all relevant windows of exposure and can be expected to reveal the longer-term effects of developmental exposures at all stages of the lifecycle.
- The SC recognises that combined exposure to multiple EASs could occur in such a way that combined toxicity could arise. The issue of combined toxicity resulting from combined exposure to multiple substances will be addressed by EFSA in a separate activity.
- The SC noted the lack of consensus in the scientific community with regard to the existence and/or relevance of low-dose effects and non-monotonic dose response curves (NMDRCs) in (eco)toxicology in connection with endocrine activity, endocrine disruption or other endpoints/modes of actions.

The SC recommends as a follow up activity to clarify in a broader context the issues of biological thresholds and criteria for adversity, combined exposure to multiple substances and NMDRCs.
The SC also underlines the need for the further development of testing strategies to generate adequate data for the identification and assessment of endocrine disrupting properties. An example has been developed in outline for fish species by the OECD.

The SC discussed several aspects that can be considered for hazard characterisation of EDs. The SC is of the opinion that hazard characterisation (e.g. establishment of a health-/ecotoxicology-based guidance value) should be based on the effect leading to the lowest health/ecotoxicology-based guidance value, irrespective of the mode of action. Such a health/ecotoxicology-based guidance value would also protect against endocrine-mediated effects occurring at higher doses. With regard to the use of severity, (ir)reversibility and potency for the hazard characterisation of EDs, the SC considers that to inform on a level of concern for endocrine disrupting substances, these elements should be evaluated in relation to the degree, duration and timing of exposure. Levels of concern are not determined exclusively by risk assessment but also by protection goals set by the risk management.

In conclusion, EDs, of natural or synthetic origin, can be identified according to three criteria: endocrine activity, adversity of effects and a plausible link between endocrine activity and adverse effect. The SC considers that a reasonably complete suite of assays is (or will soon be) available to identify and characterise the important hazards of EATS substances in mammals and fish, with fewer tests available for birds and amphibians. Furthermore, these evaluation methods should, in principle, be fit for the purpose of establishing safe doses/concentrations of EDs if (1) certain aspects (e.g. follow up of exposure in critical windows of susceptibility to later life stages) are addressed and (2) used with all available information in a weight-of-evidence approach. It should also be noted that standardised mechanistic assays for non-EATS modalities relevant to mammals, fish and other vertebrates are not yet available. For invertebrates, relevant mechanistic assays are lacking from the OECD testing suite. Finally, a range of major taxa e.g. reptiles or echinoderms, have not yet been considered by the OECD for any endocrine assay development.

Furthermore, to inform on risk and level of concern for the purpose of risk management decisions it is the opinion of the SC that risk assessment (taking into account hazard and exposure data/predictions) makes best use of available information. EDs can therefore be treated like most other substances of concern for human health and the environment, i.e. be subject to risk assessment and not only to hazard assessment.

TABLE OF CONTENTS

BACKGROUND AS PROVIDED BY THE EUROPEAN COMMISSION

The endocrine system plays a crucial role in maintaining human homeostasis and is often affected by exogenous stimuli. A range of synthetic as well as naturally occurring agents have been identified as interacting with the endocrine system. If the interaction of these exogenous substances with the endocrine system leads to adverse health effect in an intact organism or its progeny or (sub) populations, these substances are referred to as 'endocrine disruptors'[4] (EDs).

Given the range of the EU legislation under which these substances are regulated (such as plant protection products, biocides, pharmaceuticals, cosmetics, chemicals), the European Commission (EC) published its proposed Community Strategy for Endocrine Disruptors[5] in 1993. The European Parliament called on the EC in 1998 to examine many research and regulatory questions related to endocrine disruption.

The Community Strategy called for the establishment of 'a list of substances requiring priority evaluation ('ED priority list') of their role in endocrine disruption and to identify *inter alia* substances which can already be addressed under existing legislation, gaps in knowledge and specific cases of consumer use for special consideration'. On the basis of independent reviews of peer-reviewed scientific literature, and in consultation with the Commission's Scientific Committee on Toxicity, Ecotoxicity and the Environment, a candidate list of 553 synthetic chemicals and 9 hormones was published in 2000, together with a series of actions proposed to further evaluate the role of these substances in endocrine disruption. The final long-term goals of the Community Strategy are 'legislative actions' to control substances having harmful effects on humans, wildlife and/or the environment.

In recognition of the need to address the problem of endocrine disruptors, many pieces of EU legislation contain specific provisions on this issue, e.g. REACH, Food and Feed legislation, Plant Protection Products Regulation, Biocides Regulation, Regulation on cosmetics, Water Framework Directive and others. Currently, the main focus both within the EU and internationally is to agree on approaches for the identification and risk assessment of endocrine disruptors.

TERMS OF REFERENCE AS PROVIDED BY THE EUROPEAN COMMISSION

In the light of the above, EFSA is asked to advise the Commission on the following questions:

1) What scientific criteria may be used to distinguish between EDs and other groups of chemicals with different modes of action? The answer should examine the following: low-dose effects, including non-monotonic dose response, critical windows of susceptibility, threshold effects etc.

2) What scientific criteria may be used to distinguish between physiological modulation (adaptive response) and adverse effects on humans and on the ecosystem as a result of exposure to endocrine active substances?

3) Are the existing toxicity testing methods appropriate for the identification and characterisation of effects mediated by endocrine active substances (both humans and ecosystem should be considered)?

[4] "An endocrine disruptor is an exogenous substance or mixture that alters function(s) of the endocrine system and consequently causes adverse health effects in an intact organism, or its progeny, or (sub)populations.", International Programme on Chemical Safety (IPCS). 2002. Global Assessment of the State-of-the-Science of Endocrine Disruptors. WHO (World Health Organization), Geneva, Switzerland.

[5] http://europa.eu/legislation_summaries/internal_market/single_market_for_goods/chemical_products/l21277_en.htm

In developing this opinion, EFSA is requested to take account of the latest available published scientific information, including the final report 'State of the Art Assessment of Endocrine Disrupters'(Kortenkamp et al., 2011).

With a view to ensuring consistency, other Scientific Advisory Bodies, including the European Medicines Agency (EMA), the European Chemical Agency (ECHA), the European Environment Agency (EEA) and the European Commission Scientific Committees (SCCS, SCHER and SCENIHR) should be involved during the preparation of the opinion.

The Commission would ask the EFSA to provide its final opinion to the present request by March 2013.

CLARIFICATIONS TO THE TERMS OF REFERENCE

Following clarification of the problem formulation between EFSA and the EC, it was agreed that this EFSA opinion would take stock of the available information related to the three specific questions posed by the Commission in the terms of reference, namely

i) What scientific criteria should be used to identify EDs,

ii) What is an adverse effect and how can it be distinguished from physiological modulation, and

iii) Are existing toxicity testing methods appropriately covering the effects of endocrine active substances? The opinion will be based on an evaluation of existing information, current insights and scientific activities on 'endocrine disruptors', from European and other international parties.

METHODOLOGY

Composition of the working group tasked with drafting the opinion

EFSA followed its specific Standard Operating Procedure detailing the steps necessary for establishing, updating or closing a scientific working group. This procedure describes for example the process for appointing the Chair, the identification of the areas of expertise that need to be covered and the way experts with matching profiles are sought at EFSA or in the EFSA Expert Database. The areas of expertise considered to be relevant for this mandate are endocrinology (general, human and environmental), and risk assessment and toxicology (general, human and environmental). Experts with up-to-date knowledge of (OECD) test methods for endocrine disruptors were also sought.

As requested in the mandate received from the EC, other Scientific Advisory Bodies were approached to ensure consistency. This led to the identification of additional experts with a profile matching the above mentioned areas of expertise and who participate in advisory committees of the EC Scientific Committees (SCCS, SCHER, SCENIHR), EMA and EEA to join the working group in their personal capacity. In addition, representatives from the EC's Directorate-General for Health and Consumers and the Joint Research Centre – involved in the ongoing review of the EC's current strategy on EDs coordinated by the EC's Directorate-General for the Environment – , as well as EMA, EEA and ECHA were invited to participate as observers.

In addition, the EFSA Scientific Committee Unit ensured links with experts involved with international organisations working on endocrine active substances, such as OECD and WHO. A hearing expert was invited to present and discuss during the 4th meeting of the working group the WHO/UNEP draft update of the global assessment of the state of the science of endocrine disruptors.

In accordance with the Decision of the Executive Director concerning the selection of experts[6], the declarations of interests of all short-listed experts were checked for absence of conflicts of interest before they could be invited to participate in the working group to contribute in their personal capacity, as an observer or as a hearing expert. A list of the members of the working group and the observers, as well as their declarations of interest were made available on EFSA's website[7].

Sources of information

The development of this opinion was initiated by compiling important documentation previously published by various national, European and other international parties that have worked on the topic of endocrine active substances. In addition, EFSA asked its national focal points to contribute existing documents such as national position papers or reviews on endocrine active substances. In responding to the EC mandate, the SC took stock of the various (sometimes controversial) views from the various experts and fora. The SC acknowledges their informative value for the scientific issues discussed in this opinion.

EFSA did not carry out a systematic literature review on endocrine active substances or endocrine disruptors. A non-exhaustive overview of the information that was reviewed for this opinion is provided in Appendix A.

Methodologies and approaches used to evaluate the information collected

The SC and its working group applied the general principles for the evaluation of all information gathered to address the questions posed by the European Commission. These principles are described in the 2009 Guidance Document of the Scientific Committee on transparency in the scientific aspects of the risk assessments carried out by EFSA (EFSA, 2009a).

[6] Decision of the Executive Director concerning the selection of members of the Scientific Committee, Scientific Panels and external experts to assist EFSA with its scientific work (available at http://www.efsa.europa.eu/en/keydocs/docs/expertselection.pdf)

[7] See http://www.efsa.europa.eu/en/sc/scwgs.htm and https://ess.efsa.europa.eu/doi/doiweb/wg/678310

ASSESSMENT

1. Introduction

Many substances released into the environment through human activity are capable of interfering with the endocrine or hormone systems of animals and humans, which regulate the metabolism and function of the body. Such endocrine active substances (EASs) occur in a variety of chemical classes including synthetic drugs, pesticides, compounds used in industry and in consumer products, industrial by-products and pollutants, including some metals. However, one should keep in mind that there is also a large number of EASs of natural origin occurring in plants consumed as food or feed, and also some secondary metabolites from fungi that may contaminate food and feed are known to express endocrine-like activity. Examples of naturally occurring EASs are oestrogenic compounds in soy (e.g. genistein and daidzein), mycotoxins (e.g. zearalenone) in cereals, goitrogens in cabbage with the potential to inhibit iodine uptake (glucosinates), and glycirrhizine in liquorice with the potential to disturb the mineralocorticoid system.

The term 'endocrine disruptor (ED)' was first used at the Wingspread Conference in Wisconsin, USA in 1991 for those EASs, which may lead to an adverse health effect (Colborn and Corlie, 1992). WHO/IPCS developed a widely accepted definition, which is discussed later in this opinion (WHO/IPCS, 2002).

According to the above Wingspread Conference Statement on "*chemically-induced alterations in sexual development: the wildlife/human connection*", the public concern regarding EDs was originally linked to observations of reproductive and developmental toxicity in wildlife. These observations included reduced fertility, birth defects and sexual and behavioural developmental disorders. Furthermore, the public concern was also originally linked to cancer of the reproductive organs in female offspring, caused by the use of synthetic oestrogens in women during pregnancy (e.g. diethylstilbestrol (DES)). Though effects may vary between species and compounds, the Conference Statement identified four main issues arising from exposure to EDs of synthetic and natural origin: (1) they may have effects on the embryo, fetus or perinatal organism different from those on the adult organism; (2) effects are often manifested in offspring, not in the exposed parent; (3) timing of exposure in the developing organism is critical in determining future impact and character; and (4) although critical exposure may occur during embryonic development, obvious manifestations may not occur until maturity (Bern et al., 1992). A number of reports (Kortenkamp et al., 2011; EEA, 2012; WHO/UNEP, 2013) have recently analysed in detail the evidence for endocrine disruption in humans, wildlife and animal models. The reader is referred to these papers for detailed and updated information.

However, it should be noted that the above mentioned issues related to reproduction and development are not specific to EDs and that the endocrine system extends far beyond those. Furthermore, the endocrine system includes many additional signalling systems in humans and animals involving a vast number of hormonal or signalling factors, which are divided into 5 major classes: amino acid derivatives, small neuropeptides, large proteins, steroid hormones and vitamin derivatives (Jameson, 2010). In addition, numerous peptide growth factors share actions with hormones. For this reason, hormonal aspects of metabolic regulation and neurodevelopment have also recently been included in the endocrine system.

Substances that may exert their adverse effects by endocrine-related modes of action are relevant to various sectors of EFSA's activities. An EFSA technical report developed by a cross-EFSA task force was published in 2010 to clarify the state-of-play and to make recommendations for scientific and communication issues (EFSA, 2010). In that report, the term 'endocrine active substance' was used to

cover all substances that in some way may interfere with the endocrine system, but not necessarily cause adverse effects.

1.1. Scope of this opinion

The standard risk assessment paradigm consists of four steps, namely hazard identification, hazard characterisation, exposure assessment and risk characterisation, of which the latter is an integration of the first three steps (EC, 2000; WHO/IPCS, 2009). In its present mandate on EDs, the European Commission (EC) poses questions in the terms of reference related to the potential hazards (i.e. inherent properties) of a substance. This opinion therefore focuses on the first two steps of the risk assessment paradigm, namely how to identify and characterise hazards mediated by EASs.

There is a significant body of scientific evidence that some exogenous natural and synthetic chemicals can interfere with the function of animal endocrine systems (EFSA, 2010). While plant hormones are also well characterised, and there may be examples of natural or synthetic chemicals interacting with plant hormone function, the scope of this scientific opinion is the wildlife and human health effects of EDs; it does not include interference with plant hormone systems.

The first and second points of the agreed terms of reference are addressed under section 3: Criteria for identifying endocrine disruptors. The third point is discussed under section 4: Availability and appropriateness of test methods for identifying and characterising effects mediated by endocrine active substances. It is not the purpose of this opinion to develop testing strategies for EDs.

Due to the limited timeframe for the development of this opinion, a number of concepts mentioned in the original request of the ED could only be briefly discussed in this opinion:

- Thresholds aspects are considered in section 3.1, when discussing criteria to distinguish endocrine disruption from endocrine modulation.
- Low-dose effects, non-monotonic dose response curves, critical windows of susceptibility, and combined exposure to multiple substances are briefly discussed under section 4, when considering the appropriateness of test methods.
- Critical effect, consideration of severity, (ir)reversibility and potency are part of hazard characterisation (see section 5: Elements of hazard characterisation of EDs).

1.2. Overview of provisions for endocrine disruptors in EU legislation

The background information provided by the EC mentions various legislative acts that contain specific provisions on endocrine disruption. Further summaries of these are given in Appendix B. Because of the broad use of endocrine active substances, the overview of relevant legislations goes beyond EFSA's remit (risk assessment of food and feed) and covers other areas such as medicines, cosmetics, industrial chemicals and biocides. Some of the proposed concepts such as 'zero tolerance', 'negligible level of exposure', or 'levels of concern' are related to risk management and are therefore beyond the scope of this opinion.

The Scientific Committee (SC) notes that Regulation (EC) No 1107/2009 on plant protection products includes specific provisions to approve an active substance, safener or synergist, "*if it is not considered to have endocrine disrupting properties that may cause adverse effects in humans / on non-target organisms unless the exposure……is negligible*". It follows from this wording that a hazard-based approach is prescribed in the European Union for the regulation of plant protection substances exhibiting endocrine disruptive properties. A similar approach is taken for biocides in

Regulation (EC) No 528/2012. A different approach is taken in the USA and Japan, where risk assessment of all EASs is to be conducted, i.e. consideration of both hazard and exposure.

2. Definitions and terminology

A variety of inter-related terms have been used to describe the phenomenon of endocrine-mediated effects, and while these may reflect the perspectives of different stakeholders, they also have varying scientific interpretations. As an example, two such terms are used in the terms of reference for this mandate: 'endocrine disruptor', used in part i), and 'endocrine active substance', used in part ii) and iii). This section is aimed at clarifying the meaning of and relationships between these terms.

Endocrine system

The endocrine system regulates the metabolism and function of the body and is further described by WHO/IPCS (WHO/IPCS, 2002; WHO, 2012) and on the US EPA website[8]. As an inter-related system, the endocrine system influences almost every cell, organ, and function of an organism. It regulates, with the use of numerous chemical messengers, various vital functions such as metabolism, growth and development, tissue function, or mood, from conception through adulthood and into old age. This includes for example the development of the brain and nervous system, the growth and function of the reproductive system, or the regulation of blood sugar level. In order to fulfil these functions, the endocrine system uses cycles and negative feedback loops, regulating the secretion of almost all hormones. The cycles of secretion of chemical messengers, whose duration can range from hours to months, maintain physiological and homeostatic control. Once a receptor and a hormone bind, the receptor carries out the hormone's instructions by either leading to alteration of the cell's existing proteins or alteration of gene expression. Both of these actions can create reactions throughout the body. Endocrine system diseases and disorders occur when one or more of the system's components are not working well. For example, a hormonal imbalance may develop if hormones are released in too great or too small amounts, or if there are not enough receptors or binding sites.

Endocrine activity and endocrine active substance

In the scientific report on endocrine active substances (EFSA, 2010), EFSA defined an EAS as:

> *"any chemical that can interact directly or indirectly with the endocrine system, and subsequently result in an effect on the endocrine system, target organs and tissues."*

Perhaps the most familiar and well-characterised example of such an interaction is binding of a substance to a hormone receptor, e.g. the oestrogen receptor (ER). Such substances may exhibit agonist or antagonist activity (or both) in relation to the receptor, depending on the nature of its interaction with the ligand-binding site of the receptor. Other ways that a substance might exhibit endocrine activity include interference with: (i) cellular factors involved in mediating the effects of an activated hormone-receptor complex; (ii) cellular uptake of substances required for hormone synthesis; (iii) enzymes involved in hormone synthesis or metabolism/clearance; (iv) secretion of hormones from endocrine tissues; (v) binding of hormones to transport or sequestration proteins in blood plasma or within cells; and (vi) neurological function or neuro-endocrine signalling involved in regulation of endocrine function.

[8] See http://www.epa.gov/endo/pubs/edspoverview/whatare.htm

It follows from the EFSA definition of an EAS that any substance exhibiting endocrine activity, i.e. the ability to interact with one or more elements of an endocrine system, falls into the category of 'endocrine active substances'. It should be pointed out that, by definition, natural hormones (e.g. estradiol-17β, testosterone) and synthetic analogue hormones (e.g. 17α-ethinylestradiol, trenbolone) are EASs, though not all EASs are natural hormones or hormone analogues. Moreover, the scope of the definition of the endocrine system determines the variety of substances that may be identified as EASs; this variety is contingent on our increasing understanding of the endocrine system.

Physiological modulation via the endocrine system

By 'physiological modulation' which is mentioned in the terms of reference, the Scientific Committee understands 'physiological modulation via the endocrine system'; the terminology 'endocrine modulation' is therefore used throughout the rest of this document.

As noted in the Final Report on the State of the Art Assessment of Endocrine Disrupters (SAAED) the compensatory feedback mechanisms that typify endocrine systems provide homeostatic capacity, which is adaptive (Kortenkamp et al., 2011). Exposure to exogenous substances that exhibit endocrine activity (i.e. EASs) may stimulate modulation in these feedback systems. If this modulation and its effects are temporary and within the homeostatic capacity of the endocrine system of the exposed organism, the effect of the substance might be considered endocrine modulation and hence non-adverse (see below).

Endocrine disruption and endocrine disruptor

The SAAED points out that regulation of chemicals on grounds of their toxicological properties "*cannot proceed without finding scientifically sound definitions of the effects in question (here: endocrine disruption*" (Kortenkamp et al., 2011). Consequently, there have been several widely cited attempts at defining an ED.

In 1996, the US EPA proposed the following definition of an ED during a workshop (Kavlock et al., 1996):

> "*An ED is an exogenous agent that interferes with the synthesis, secretion, transport, binding, action or elimination of natural hormones in the body, which are responsible for the maintenance of homeostasis, reproduction, development and/or behaviour*."

Whilst the US EPA was the first agency to define endocrine disruption as a mode of action, this definition is ambiguous in not differentiating adequately between compensatory/homeostatic changes (i.e. endocrine modulation as discussed above) and those that lead to adverse health effects (see discussion on adversity below).

The Weybridge definition (EC, 1997) of an endocrine disruptor makes explicit reference to adversity:

> *"An ED is an exogenous substance that causes adverse health effects in an intact organism, or its progeny, secondary to changes in endocrine function."*

as does the WHO/IPCS definition (WHO/IPCS, 2002):

> *"An endocrine disrupter is an exogenous substance or mixture that alters function(s) of the endocrine system and consequently causes adverse health effects in an intact organism, or its progeny, or (sub)populations."*

The WHO/IPCS definition departs from the Weybridge definition in two main respects: 1) use of 'consequent' in place of 'secondary', which has been interpreted as placing a great onus on

demonstrating cause-effect linkage between endocrine activity and an adverse health effect; 2) adding (sub)populations, which may be seen to make the definition more directly applicable to ecotoxicology (see below). A survey presented in the SAAED Final report showed that most EU Member States acknowledge the WHO/IPCS definition, and this definition is extensively discussed in SAAED (Kortenkamp et al., 2011).

Recently, the Endocrine Society published a statement of principles on endocrine disruptors and public health protection (Zoeller et al., 2012), in which another definition of an ED has been proposed:

> *"An ED is an exogenous chemical, or mixture of chemicals, that interferes with any aspect of hormone action."*

This definition is a simplification of the definition developed by the US EPA (Kavlock et al., 1996). Zoeller et al. (2012) emphasise that while it is critical for hazard identification to be able to capture the sensitivity of human and wildlife to chemicals that pose a potential risk, the ability of a chemical to interfere with hormone action (i.e. the hazard), is of itself a reliable predictor for adverse outcomes. In their view, uncertainty in the relation between the endocrine activity and the manifestation of an adverse consequence relate to the dose, duration and timing of exposure (Zoeller et al., 2012). The above-mentioned areas of uncertainty would normally be addressed in a risk assessment that considers both hazard and exposure. The Scientific Committee stresses that a positive signal for endocrine activity obtained from *in vitro* or *in vivo* testing does not automatically imply that an adverse effect will be observed in an intact organism. As such, the SC cannot agree with the above views, as the Zoeller et al. statement implies that an ED is equivalent to an EAS as defined by EFSA. The SC however concurs with the suggestion from the authors that the uncertainties associated with the hazard-based approach for the management of an ED should be addressed using a risk assessment approach, i.e. considering both hazard as determined *in vivo* and exposure (see section 5).

In the OECD Guidance Document on Standardised Test Guidelines for Evaluating Chemicals for Endocrine Disruption (OECD, 2012a), an operational definition of a 'possible endocrine disruptor' is provided:

> *"to mean a chemical that is able to alter the functioning of the endocrine system but for which information about possible adverse consequences of that alteration in an intact organism is uncertain."*

This term and its definition were intentionally chosen to be distinct from 'potential endocrine disruptor', so it could be unambiguously applied to chemicals being tested in the Conceptual Framework (OECD, 2012a) for purposes of confirmation or elaboration of suspected endocrine activity (i.e. resolving hazard identification and hazard characterisation).

WHO/IPCS (WHO/IPCS, 2002) defines a 'potential endocrine disruptor' as follows:

> *"A potential endocrine disruptor is an exogenous substance or mixture that possesses properties that might be expected to lead to endocrine disruption in an intact organism, or its progeny, or (sub)populations*."

Presumably those properties represent the innate ability of the substance (e.g. as might be identified from *in vitro* evidence) to interact with elements of the endocrine system of an exposed organism, and a 'potential endocrine disruptor' could therefore be considered equivalent to an EAS. It follows that a potential endocrine disruptor is a substance for which uncertainty exists about the realisation of the adverse effect in an intact organism (i.e. *in vivo*).

The SC considers that the meaning of possible/potential ED overlaps with the definition of an EAS, as suggested previously by EFSA (EFSA, 2010). For these terminologies, EAS is used in this opinion.

Adversity

The US EPA definition (Kavlock et al., 1996) and the simplified version proposed by the Endocrine Society, define an ED in relation to mode of action alone, which is comparable to the definition of an EAS (as discussed above). In contrast, Weybridge and WHO/IPCS define EDs in terms of both mode of action and adversity of the effect.

An adverse effect has been defined by WHO/IPCS (2009) as follows:

> *"Change in the morphology, physiology, growth, development, reproduction, or, life span of an organism, system, or (sub)population that results in an impairment of functional capacity, an impairment of the capacity to compensate for additional stress, or an increase in susceptibility to other influences."*

The later part of this definition *"an impairment of the capacity to compensate for additional stress, or an increase in susceptibility to other influences"* is context- dependent, i.e. there is a reduced capacity of *a specific population sample of organisms*; As such, these organisms are responding physiologically to their environment: the testing laboratory, micro/mesocosm, or the wild. Therefore, by this definition, adversity cannot be presumed or established from the hazard alone. This is echoed in the National Research Council of the US National Academies (NRC) report on toxicity testing in the 21st century (NRC, 2007), prepared for the US EPA, which states:

> *"The consequences of a biologic perturbation depend on its magnitude, which is related to the dose, the timing and duration of the perturbation, and the susceptibility of the host. Accordingly, at low doses, many biologic systems may function normally within their homeostatic limits. At somewhat higher doses, clear biologic responses occur. They may be successfully handled by adaptation, although some susceptible people may respond. More intense or persistent perturbations may overwhelm the capacity of the system to adapt and lead to tissue injury and possible adverse health effects."*

Whether or not adversity results from exposure of the intact organism or (sub)population to an EAS (as a hazard) is determined by the nature (strength) of the relationship between the hazard and the biological response (effect), the state of the organism/(sub)population (the other stresses it faces) and the intensity of the exposure to the hazard. Some of this information can be gathered through hazard characterisation.

As noted above with respect to definitions, the WHO/IPCS definition departs from the Weybridge definition in adding '(sub)populations', which may be seen to make the definition more directly applicable to environmental hazard and risk assessment of chemicals. For the environment, adversity is seen in the context of the presumed protection goal, generally considered to be population stability (i.e. effects on individuals may be acceptable if they are not expected to have implications for the population). It should be noted that the difference in level of biological organisation at which the adversity (population) and endocrine mode of action (individual) are determined, presents something of a challenge in applying definitions of endocrine disruption for regulatory purposes in the ecotoxicological context. An EAS should therefore only be identified as an environmental ED if it can be demonstrated or plausibly argued, that a wildlife[9] population is likely to be affected. Evidence that can be used for this includes laboratory or field data on such endpoints as growth, development and reproductive success, which in population models or other predictive methods lead to an extrapolation of biologically significant effects on population size or stability.

[9] The term 'wildlife' as used herein, covers non-target species only and does not cover wildlife intended to be controlled by the application of regulated products (i.e. target species). It is also noted that whilst the considerations in this opinion do not specifically address other animal populations (e.g. farmed animals or companion animals), the same principles could be applied to these groups of animals, taking account of potentially additional specific protection goals.

- Conclusion on the definitions that will be used by EFSA

The EFSA Scientific Committee concludes that the WHO/IPCS (2002) definition of an ED (which is the same definition as the one used by the EC in the background provided to the terms of reference) and the WHO/IPCS (2009) definition of an adverse effect should be endorsed as working definitions in this opinion.

It should be noted that in line with protection goals embedded in EU legislation, adverse effects are addressed at the level of the individual(s) for human health and at the level of the (sub)population for wildlife.

An EAS is a substance that has the ability to interact directly or indirectly with the endocrine system, and subsequently results in an effect on the endocrine system, target organs and tissues; there is however uncertainty as to whether it is likely to produce adverse effects measured on apical endpoints *in vivo*. In some cases, EASs may only produce biological changes that lie within an organism's homeostatic capacity, or be detoxified by metabolism, and would therefore not be expected to cause adverse effects in the intact organism. In other cases, EASs may perturb homeostatic systems and thus cause adverse health effects at the level of the whole organism. The latter types of EASs can be termed EDs.

3. Criteria for identifying endocrine disruptors

This section addresses the first and second parts of the agreed terms of reference: 1) what scientific criteria should be used to identify EDs and 2) what is an adverse effect and how can it be distinguished from endocrine modulation.

Following the decision of the EFSA Scientific Committee to use the WHO/IPCS definition of ED (see section 2) as a working definition for this opinion, an ED is defined by three criteria: i) the presence of an adverse effect in an intact organism or (sub)population; ii) the presence of an endocrine activity; and iii) a plausible or demonstrated causal relationship between the endocrine activity and the adverse effect.

This chapter will only focus on issues that are specific to endocrine disruption. As such, the SC underlines that the question of the difference between physiological modulation (e.g. endocrine modulation) and adversity is not unique to EDs and should therefore be addressed in a broader context (see section on recommendations).

3.1. Adversity and criteria to distinguish endocrine disruption from endocrine modulation

As described in section 2, the SC decided to use the definition for adverse effect produced by WHO/IPCS in 2009:

> *"Change in the morphology, physiology, growth, development, reproduction, or, life span of an organism, system, or (sub)population that results in an impairment of functional capacity, an impairment of the capacity to compensate for additional stress, or an increase in susceptibility to other influences."*

This is a generic definition of adversity, which is not specific to the endocrine system.

The definition of an ED agreed upon for this opinion (see section 2) implies that the adverse effect should be observed in an intact organism (e.g. not in *in vitro* systems, or castrated or ovariectomised test animals). In relation to the environment, the SC reiterates that the focus of any ecotoxicological hazard assessment is the protection of non-target populations. Therefore, it is important to establish that the adverse effects observed in experimental test animals are relevant to populations. Particular attention should hence be given to the adverse consequences on reproduction, growth/development, disease incidence and survival in one or more environmental species as these are the effects most likely to impact on population recruitment and stability.

As noted in SAAED, the compensatory feedback mechanisms that typify endocrine systems provide homeostatic capacity against various endocrine perturbations. Exposure to exogenous substances that exhibit endocrine activity may stimulate modulation in these feedback systems. If this modulation is temporary and/or within the homeostatic capacity of the endocrine system of the exposed organism, the effect of the substance might be considered 'endocrine modulation'. Alternatively, if the body is unable to compensate for the induced changes within its limits of homeostasis (e.g. for some endocrine modalities/axes/pathways during critical periods of development or at high doses), the threshold of adversity is crossed, and the observed changes are to be considered as adverse. It should be noted that a transient endocrine modulation, which is simply adaptive in the adult organism, can result in permanent adverse changes in the developing organism, as the system responsible for the normal endocrine homeostasis of the latter may not be fully developed yet (WHO, 2012).

Overall, therefore, endocrine effects become adverse either by elicitation of a sub- (or supra-) normal response or persistence in a physiological state that is intended to be transitory. The SC notes that interpretation of endocrine disrupting effects should also consider nutritional status, as the latter may influence the former. This is well-known for thyroid disruption, where the effect strongly depends on the iodine status of the individual.

The SC is in agreement with Kortenkamp et al. (2011) that, since points have not been defined where 'threshold of adversity' is crossed, it is difficult to propose specific criteria to differentiate between effects that represent an endocrine modulation and adverse effects on the endocrine system. Expert judgement will therefore be required to assess on a case-by-case basis the toxicological relevance of such changes. In general, transient, inconsistent and minor fluctuations at the biochemical and molecular level may be considered adaptive (i.e. non-adverse), whilst sustained, consistent and permanent changes at the cell-, organ- or organism-level, resulting in pathology or functional impairment *in vivo*, as well as altered timing of development, may be considered adverse.

The point at which endocrine modulation becomes an adverse effect cannot be determined on the basis of an absolute response value, but on the basis of a relative response (compared to the control/background response). The SC is therefore of the opinion that, as adversity is a prerequisite for identifying a substance as an ED, it is necessary to determine a biological threshold between endocrine modulation and adverse effect. For the time being, it is difficult to propose generic criteria to determine when this biological threshold is crossed. This is therefore likely to be done on a case-by-case basis through expert judgement.

For most toxic processes, it is generally assumed that there is a threshold of exposure below which no biologically significant effect will be induced (Dybing et al., 2002; WHO/IPCS, 2009). The existence of dose thresholds cannot be proven or ruled out by experimental approaches, because all methods for measuring effects have their limits of detection which will obscure thresholds, if they exist (Kortenkamp et al., 2011). However, the presence of homeostatic and cytoprotective mechanisms, and the redundancy of cellular targets, mean that a certain degree of interaction of the substance with the critical sites or their occupancy must be reached in order to elicit a toxicologically relevant effect (Dybing et al., 2002). Below this critical (threshold) level of interaction, homeostatic mechanisms would be able to counteract any perturbation produced by xenobiotic exposure, and no structural or

functional changes would be observed. In certain developmental stages, homeostatic capacity is limited and this will affect the sensitivity of the organism.

Specificity - Endocrine effects secondary to non-endocrine related toxicity
Adverse effects occurring in the presence of marked toxicity tend to represent the unspecific and generalised response of the body to the chemical insult (e.g. arising from the saturation of kinetic processes, through overwhelming of defence/repair mechanisms or through elicitation of stress responses). In the presence of generalised toxicity, it is most likely that the endocrine system will also be affected, especially considering that one of the main functions of the endocrine system is maintenance of homeostasis in response to different stimuli/stressors. The SC therefore concludes that endocrine-related adverse effects secondary to marked toxicity caused by a non-endocrine mode of action should not be considered as an endocrine-mediated mode of action.

3.2. Endocrine activity

As noted in section 2, 'endocrine activity' implies the inherent ability of a substance to interact or interfere with one or more components of an endocrine system and 'endocrine active substance' indicates a substance that is able to interfere with the endocrine system, but does not necessarily lead to adverse effects. Therefore, the Scientific Committee considers endocrine activity as a collection of modes of action, potentially leading to adverse outcomes, rather than as a (eco)toxicological hazard in itself.

Endocrine activity information could be obtained from existing information, read-across, *in silico* tools, *in vitro* and *in vivo* screening assays (Levels 1, 2 and 3) of the current OECD Conceptual Framework[10] (OECD, 2012a) or from other mechanistic investigations. These mechanistic studies should be evaluated on their merits on a case-by-case basis.

3.3. Causal relationship between endocrine activity and adversity

In order to conclude that a substance is an ED there must be a reasonable evidence base for a biologically plausible causal relationship between the induced endocrine perturbation/activity and the adverse effects seen in intact organism studies. The SC underlines that there should be no difference between the level of evidence needed to demonstrate the endocrine activity/mode of action and the level of evidence needed to demonstrate the adverse effect.

As for any other (eco)toxicological hazard, endocrine-mediated adverse effects may be identified in standard toxicological tests that are routinely performed to fulfil the requirements of various regulatory programmes. In particular, endocrine-mediated toxicity may be detected in repeated-dose, reproductive and developmental toxicity, and carcinogenicity studies, although supplementary and more focussed studies, such as mechanistic studies investigating the potential for endocrine activity, may be necessary to decide whether a causal relationship between the observed adverse effects and an endocrine activity is biologically plausible. The SC is aware that recent reports (EEA, 2012; WHO/UNEP, 2013) question the adequacy of these tests, with respect to combined exposures of EASs (see further section 4.7.2.2).

[10] See: http://www.oecd.org/env/chemicalsafetyandbiosafety/testingofchemicals/oecdconceptualframeworkforthetestingandassessmentofendocrinedisruptingchemicals.htm

Demonstration of all the key events of an endocrine mode of action leading to the adverse outcome is not necessary, as this requires a very high burden of proof. However, it is important that there is logical and plausible reasoning to explain any (potential) causal relationship between the observed endocrine activity and the endocrine-mediated adverse effects. This concept of 'plausibility' implies expert judgement. A minimum set of criteria was described by Bradford Hill (Bradford Hill, 1965) to provide adequate evidence of a causal relationship between an incidence and a consequence (e.g. exposure and ill health), including biological plausibility, consistency of findings, specificity, predictivity, coherence, concordance of dose response relationships and temporal associations and characterisation of uncertainties.

The SC considers that a test does not necessarily need to be standardised in order to demonstrate either the endocrine activity or the adverse effect. Any data whose robustness has been demonstrated is acceptable for the hazard assessment. This assessment of the individual data for robustness needs to include a judgment on: 1) The validity of the method or model used (i.e. Is the method/model sufficiently predictive?); 2) The adequacy of the individual pieces of information, composed of the elements reliability (i.e. Was the method/model applied correctly?); and 3) Its relevance (i.e. Was the method/model appropriate for the intended purpose?).

Weight-of-evidence

All the available information on adversity and endocrine activity (*in silico*, *in vitro* and *in vivo* data including observational studies) should be considered together, by adopting a weight-of-evidence approach. Guidance on how to consider weight-of-evidence has been provided by WHO (e.g. Boobis et al., 2006; Boobis et al., 2008) and SCENIHR (2012). Evidence for adversity and endocrine activity should therefore be evaluated in parallel and not in sequence.

In relation to human health, the default assumption of any adverse effect seen in toxicity studies is that the effect is relevant to humans. This assumption can be rebutted with sound scientific data showing non-relevance. It is proposed that a structured framework, e.g. the WHO/IPCS human relevance framework (e.g. Boobis et al., 2008), is used to analyse the available evidence and biological plausibility to facilitate a robust and transparent conclusion.

As discussed in a previous opinion of the Scientific Committee (EFSA Scientific Committee, 2011), "*Absence of Evidence is not Evidence of Absence*". It follows that it is not possible to prove the absence of any effect (including endocrine-related effects) and therefore, a negative test would not allow the exclusion of endocrine activity. The various types of above-described data should therefore be considered as part of a weight-of-evidence approach to conclude about whether a substance is an ED. The SC underlines the need for a testing strategy (see section 4.8) to generate appropriate data.

4. Availability and appropriateness of test methods for identifying and characterising effects mediated by endocrine active substances

This section addresses the third part of the agreed terms of reference: are existing toxicity testing methods appropriately covering the effects of endocrine active substances? This section therefore reviews available methods that can help to identify an endocrine activity (both human health and the environment are considered). This is termed mechanistic information.

Furthermore, since the definition of an ED endorsed by the SC includes the criterion of an adverse effect, this section also reviews available methods that can help to identify and characterise endocrine-mediated adverse effects. Various endocrine-sensitive endpoints are generally included in standard

toxicity tests. These tests for apical endpoints are designed to identify and characterise the *in vivo* adverse effects that may be endocrine-related.

Utilising the OECD revised Conceptual Framework (CF) (OECD, 2012a) as the starting point for assessing the adequacy of the available (internationally) standardised / validated[11] test methods for the identification and characterisation of effects mediated by EASs as well as the analyses and recommendations from SAAED (Kortenkamp et al., 2011), this section addresses the current human health and environmental test methods available from the OECD and US EPA websites[12], and related guidance focusing particularly on validated tests, test guidelines (TGs) and the relevant test projects in the OECD work plan (see also Appendix C).

During this stock-taking process, potential gaps in the current testing suite of internationally validated methods and sufficiently standardised protocols close to validation were reported. It is outside the scope of this section to consider test methods that were not internationally standardised for regulatory toxicity testing but used predominantly for academic research purposes.

4.1. Background to OECD guidance on test methods for endocrine active substances

In 1998, the OECD initiated a high-priority programme to revise existing and to develop new TGs for the screening and testing of EDs. Since then, a number of potential assays have been developed into TGs and others are in development. The screens and tests are summarised within the OECD revised CF for Testing and Assessment of Endocrine Disrupters (OECD, 2012a). The CF schematically lists and provides a guide to the OECD TGs and standardised test methods available, under development or proposed that can be used to evaluate chemicals for endocrine disruption. Substantial guidance on the use and interpretation of the tests and on the identification of potential information gaps is provided in OECD Guidance Document (GD) 150.

All the validated TGs have been approved by all OECD Member Countries (i.e. they are consensus documents). The SC therefore acknowledges the relevance and utility of the test methods and appreciates the ongoing validation review work and TG development to address further endocrine relevant endpoints.

For the development of the following sections, the SC also considered the relevant OECD Detailed Review Papers (DRP) such as DRP No 178 (OECD, 2012b) on novel methods and endpoints for evaluating endocrine disruptors, DRP No 97 (OECD, 2008) on metabolism for endocrine disruptor tests, DRP No 57 (OECD, 2006) on the Thyroid Hormone Disruption Assays, and DRP No 135 (OECD, 2010) on Environmental Endocrine Disruptor Screening: the Use of Oestrogen and Androgen Receptor Binding and Transactivation Assays in Fish. The SC also appreciates that the OECD ED-related expert working groups and validation management groups are currently developing many of the recommendations contained within these DRPs, and that this information is not yet publicly available, as it has not been finalised.

In 2009, hosted by Denmark, the OECD held a workshop on *"OECD Countries' Activities Regarding Testing, Assessment and Management of Endocrine Disrupters"*. This workshop recommended further work for OECD, in particular:

[11] 'Internationally standardised' means that these assays have been validated and approved for test guidelines (TGs) use by the OECD and the mutual acceptance of data (MAD) principle applies. 'Validation' refers to the formal process through which assays are shown to dependably work as intended and thus produce reliable results that can be compared across studies and laboratories (Hartung et al., 2004; OECD, 2005a). Whilst lack of validation does not necessarily mean an assay is invalid, it does mean results from the assay could be considered unreliable by regulatory agencies, depending on a number of other factors, such as inter-laboratory reproducibility.

[12] http://www.oecd-ilibrary.org/content/package/chem_guide_pkg-en; http://www.ehso.com/testmethodsdl.php

(i) The development of a guidance document for the assessment of endocrine disruptors (now published as OECD GD 150), and

(ii) The revision of the 2002 Conceptual Framework for Testing and Assessment of Endocrine Disrupters:

> The GD on Standardized Test Guidelines for Evaluating Chemicals for Endocrine Disruption, GD 150 (OECD, 2012a), produced by the Endocrine Disrupters Testing and Assessment Task Force Advisory Group (EDTA-AG) was co-authored by a large group of international experts and endorsed by OECD member countries. It provides guidance on how tools from lower levels of the CF can be used to determine which higher level tests are needed for a specific chemical, to increase evidence that it is or is not an ED. The EDTA-AG selected three chemical case studies to take through the revised CF to demonstrate the value of data interpretation pathways (OECD, 2012c).
>
> The revised CF is contained within GD 150, and the guidance on data interpretation covers 27 of the assays listed in the CF, including *in vitro*, mammalian and non-mammalian methods. Methods currently in (pre)validation are addressed in an annex. The GD is intended to be a living document which will be updated periodically in the light of experience, and as new assays are validated and developed into approved TGs.

(iii) The development of a DRP on endpoints not included in existing test guidelines (see OECD information on OECD Work Related to Endocrine Disrupters (2012)), which is now published:

> OECD Detailed Review Paper No 178. The State of the Science on Novel *In vitro* and *In vivo* Screening and Testing Methods and Endpoints for Evaluating Endocrine Disruptors (OECD, 2012b).

Introduction to the OECD revised Conceptual Framework

The information contained within the CF addresses the complexity and comprehensive relevance of information available in five ascending Levels: from Level 1 (existing information and non-test information, which should guide the initial needs for testing and assessment), to Level 2 (selected *in vitro* endocrine mechanistic/mode of action test methods), to *in vivo* selective endocrine mechanistic screening methods in Level 3, *in vivo* apical tests (for adverse effects) which include endocrine relevant endpoints in Level 4, and more comprehensive data over more extensive parts of the life cycle in Level 5.

The mechanistic studies in Levels 2 and 3 of the CF can identify endocrine activity as a prerequisite for a substance to be considered an EAS, but individually are often insufficient for ED identification. The output from Level 2 and 3 tests can be utilised for subsequent prioritisation for further testing, using the methods with apical endpoints (e.g. those of Levels 4 and 5). The latter identify the adverse effect that is a prerequisite for an EAS to be regarded an ED. The GD is based on the need to evaluate data on EDs using a weight-of-evidence (WoE) approach, priority setting and adverse outcome pathway (AOP) approach (NRC, 2007; Ankley et al., 2010; OECD, 2012e). In principle, no single assay is likely to provide all the information needed to decide whether a substance is an ED (according to the WHO/IPCS definition endorsed by the SC) because of the need to provide both mechanistic information showing how the substance interacts with the endocrine system, and apical information describing the adverse effects this interaction may cause. The results from a combination of tests increase the WoE and further elucidate the AOP. As detailed in the OECD CF (OECD, 2012a), it is possible to enter and exit the CF at all Levels, depending upon the nature of the existing information, and the regulatory needs for testing and assessment. The assessment of each substance should be made

on a case-by-case basis, again, taking into account all the available information. There may be some occasional instances, whereby a pattern of responses characterised by a spectrum of adverse effects *in vivo* clearly known to be caused by an established endocrine activity (as identified from existing (Level 1) information) may be present. For example, male rats exposed to anti-androgenic substances may exhibit a range of symptoms such as non-descended testis or external genitalia malformations (Foster, 2005). In such cases, the adverse effects could be regarded as being diagnostic of endocrine disruption, and endocrine activity could be inferred rather than being demonstrated.

See Appendix C for an extract of the CF list and sections 4.4, 4.5 and 4.6 for further details on the tests contained within each Level of the CF for mammals, aquatic organisms and birds. For each of these tests, the corresponding test guideline presents and discusses both the advantages and limitations of the test and can be consulted directly.

4.2. OECD revised Conceptual Framework Level 1: existing data and non-test information

The OECD GD (OECD, 2012a) states that "*It is important to emphasise that before conducting any assessment of data from an endocrine disruption screen or test, all existing information on the test chemical should be collated*". Therefore Level 1 of the CF consists of: Physical & chemical properties, e.g., Molecular Weight reactivity; volatility, biodegradability; all available information, including epidemiological and field studies; (eco)toxicological data from standardised or non-standardised tests; read-across; chemical categories; (Quantitative) Structure Activity Relationship ((Q)SARs) and other *in silico* predictions; and Absorption, Distribution, Metabolism and Excretion (ADME) model predictions.

4.2.1. Considerations for the use of epidemiological, field and experimental animal information when characterising the hazards for humans or wildlife of exposure to endocrine active substances

There are important considerations that first need to be taken into account when both designing and interpreting epidemiological studies to assess plausible causal relationships in relation to EASs. These include the following confounding factors: 1) environmental EASs are numerous and ubiquitous, 2) humans are generally exposed to low levels and to multiple substances and, 3) other 'environmental' conditions such as lifestyle factors, including nutritional status, may come into play. Precise estimates of the exposure (by using appropriate biomarkers, where available) to a particular EAS together with the identification of the related critical developmental stage(s) are major challenges in studies aimed at assessing the effects of EASs in humans. Furthermore, reproductive endpoints in humans may be characterised by a broad range of natural variability in a heterogenous population. Therefore, their assessment in epidemiological studies is challenging.

For field data, in principle, there is no major difference between the evaluation of effects of EDs and non-EDs. However, in view of the need to be confident whether an adverse effect of an ED is likely to have consequences at the population level, the use of field data, if available, may be valuable. In the absence of such data, regulators must be confident of being able to extrapolate from laboratory data on endpoints such as growth and reproduction to potential effects on populations, ideally but not necessarily through the use of population modelling. It is acknowledged that some effects, particularly those affecting individual behaviour, may not easily be apparent when observing at (sub)population level.

Animal studies have contributed considerably to our understanding of health disorders resulting from endocrine disruption in humans and wildlife, although many of these studies have been carried out with high exposure levels that are not necessarily relevant to humans or wildlife. For regulatory purposes, whilst maintaining a sufficient number of doses to obtain a robust dose response curve, experiments with animals should also be made at exposure levels relevant for humans or wildlife.

Considering the uncertainties surrounding the effects of EASs on human health or wildlife and the limitations of extrapolation from non-clinical data, it can be concluded that the conduct of the above types of studies, in spite of inherent challenges such as methodological set-up, quality or dose/concentrations used, remains an essential component of the evaluation of possible effects of EASs in large populations. Further descriptions of these types of studies are provided in Appendix D, highlighting critical issues which characterise the conduct and most importantly the interpretation of such studies.

4.2.2. Computational toxicology and non-test methods for screening substances for endocrine activity

4.2.2.1. Background

The conventional methods of testing chemical toxicity require the use of animals. The Scientific Committee and the Panel on Plant Protection Products and their Residues recognised in previous opinions (EFSA, 2009b; EFSA Panel on Plant Protection Products and their Residues, 2012a), the importance of minimising the use of experimental animals during food/feed risk assessments, as well as screening substances for priority setting and subsequent testing. In this regard, computational (*in silico*) toxicology offers a non-animal test approach based upon biological and chemical properties, for rapid screening of chemicals against some toxicological endpoints related to endocrine activity. These tools belong to Level 1 of the OECD CF for ED testing.

In silico and non-animal test method approaches include molecular modelling tools; structure activity relationships (SAR) upon which physico-chemical read-across and chemical categories are based; predictive statistical models such as (Quantitative) Structure Activity relationships ((Q)SARs); databases; and expert systems.

These tools may be freely or commercially available; however, the focus of this section is on the freely available data and tools. Additional information on the commercially available tools (such as Topkat, Leadscope, Derek, Meteor, Hazard Expert) can be found in the reviews cited herein. It is important to note that the regulatory acceptance of these tools is currently very limited. However, the tools may be useful for screening purposes or when designing an integrated testing strategy. Extensive efforts are underway at European and international levels to improve the quality, reliability, use and integration of such tools.

The following sections provide an overview of the different tools and their relative endpoint specific utility.

4.2.2.2. *In silico* 3-Dimensional Molecular Modelling Approaches

In silico 3-Dimensional (3D) molecular modelling tools based upon the receptor/enzyme-ligand docking simulation in protein crystal structures and/or homology modelling of receptors have been reliably used in pharmacological research and development for decades. There are numerous published examples in the medicinal and pharmacological literature. Molecular dynamics are useful for revealing facets of activation and inactivation, so improving mechanistic understanding and

predicting molecular ligand binding activity. These 3D computational tools can have a high level of accuracy, and in the last 10 years have been explored and applied in the fields of EASs chemical risk assessment (e.g. Jacobs et al., 2003; Jacobs, 2004; Akahori et al., 2005; D'Ursi et al., 2005; Vedani et al., 2012).

4.2.2.3. Chemical categories and read-across

As defined by the OECD, 'a chemical category is a group of chemicals whose physicochemical and human health and/or ecotoxicological properties and/or environmental fate properties are likely to be similar or follow a regular pattern, usually as a result of structural similarity' (OECD, 2007a, 2009a). In practice, endpoint information for one chemical (the source chemical) is used to predict the same endpoint for another chemical (the target chemical), which is considered to be 'similar' in some way (usually on the basis of structural similarity or on the basis of the same mode or mechanisms of action). However, there are a number of limitations that should be kept in mind including the fact that molecules of similar structure sometimes have dissimilar biological activity (Van Drie, 2003).

The freely available OECD (Q)SAR Toolbox (see also section 4.2.2.5) provides a platform for grouping chemicals that share structural and/or functional activity similarities, using a substantial set of high quality databases. The strategy for screening in the Toolbox involves grouping and searching the available databases using the structure alerts from the target substance, e.g. for ER binding potential (OECD, 2007a, 2009a).

4.2.2.4. Predictive (Q)SAR models

SAR is an approach designed to discover relationships between chemical structure (or structural-related properties) and biological activity (or target property) of studied compounds. (Q)SARs are algorithm–driven models based on a (quantitative) structure-activity relationship. They provide mathematical descriptions of the biological activity of a group of chemical substances in terms of one or more of their physicochemical properties. The biological activity, including toxicity, can occur at the molecular level, the tissue, target organ, biological system and whole organism level. (Q)SARs can be local - in that the compounds used to derive the model are all closely related (i.e. are from a 'narrow' applicability domain) either structurally, by mode of action, or both, - and these are usually more robust. However, they are not applicable to the wider chemical universe. Global (Q)SARs use a more structurally diverse chemical training set; so they are usually applicable to a wider 'domain' but are generally less robust in their predictivity. The principles for the development of reliable and robust (Q)SAR models have been agreed by the OECD (OECD, 2004). In brief, the models need to be relevant to specific regulatory endpoints, to have been developed using a transparent methodology and an unambiguous algorithm, and tested rigorously for robustness and predictivity against external datasets, i.e. those not used in modelling. The 'applicability domain' for each model is also determined and clearly defined.

A fully tested (Q)SAR model generally yields reliable estimates of toxicity of untested chemicals (70% or higher accuracy) as long as they fall within the model's prediction domain. Thus, a (Q)SAR model will have a limited value for assessment of those chemicals that are outside its prediction space. Combining assessments from more than one (Q)SAR model with additional information from, e.g. structural alerts and read-across estimates from analogous molecules in a WoE approach can improve the utility of these tools and the reliability of the overall *in silico* assessment.

The available (Q)SAR models relevant to EASs have been reviewed, amongst others by Benfenati et al. (2005), Lo Piparo and Worth (2010), and Castello and Worth (2011) and by external contracts

commissioned to assist the development of an EFSA opinion on the toxicological relevance of pesticide metabolites (EFSA Panel on Plant Protection Products and their Residues, 2012a).

Since the most widely studied mechanism of action for endocrine disruption has been ER binding, it is not surprising that the first robust EASs relevant (Q)SARs to emerge have been for that modality. The Danish National Food Institute has published environmental chemical relevant models on the ER, androgen receptor (AR) (Jensen et al., 2008; Vinggaard et al., 2008) and the Pregnane X Receptor (PXR) (Dybdahl et al., 2012). Combinations of 3D modelling tools with (Q)SARs such as 3D (Q)SAR models are also available for steroid hormone receptors such as the AR, glucocorticoid receptor (GR), PXR, peroxisome proliferator-activated receptor (PPAR) gamma, progesterone receptor, and thyroid receptors (TRs), and the cytosolic aryl hydrocarbon receptor (AhR) (Jacobs, 2004; Vedani et al., 2012). There are models published for other endocrine modalities, such as the retinoid X receptor and oestradiol sulfotransferase, but only a few have so far been validated or are sufficiently broad to be of general use. The large diversity of chemicals that might interfere with the endocrine system and the many potential molecular targets make the establishment of a single (Q)SAR for EASs an unrealistic goal (Jacobs et al., 2008; OECD, 2008). (Q)SAR models for the more complex endpoints, such as developmental and reproductive toxicity, are not considered reliable on their own (Maslankiewicz et al., 2005; Lo Piparo and Worth, 2010; JRC, 2011; Worth et al., 2011; EFSA Panel on Plant Protection Products and their Residues, 2012a), but again may have WoE utility if used appropriately in combination with other *in silico* models and tools.

4.2.2.5. Databases and freely available software platforms

Databases of EASs are available and include the US FDA Endocrine Disruptor Knowledge Base (EDKB); the Endocrine Active Substances Information System (EASIS); the US EPA Endocrine Disruptor Screening Program (EDSP) Universe of Chemicals. In the EU, the Endocrine Active Substances Information System (EASIS)[13] is under development to update the existing database of the EC's DG Environment.

Freely available software platforms

A major software platform that is useful for *in silico* screening of EASs is the OECD (Q)SAR Toolbox. This is a software application intended to fill gaps in (eco)toxicity data needed for assessing chemical hazards; it incorporates databases on physico-chemical data and on experimental (eco)toxicological data, a range of (Q)SAR models of varying quality, together with (Q)SAR modelling tools and Expert Systems. It incorporates a large number of databases for various toxicological endpoints; however, to date the only endocrine-specific database in the Toolbox is the OASIS ER Binding Database that consists of 1460 diverse compounds with Relative ER Binding Affinity data, where the data generated is all relative to the positive control 17-beta-oestradiol.

The (Q)SAR models contained in the Toolbox include the MultiCASE RBA (Q)SAR, which is based on a hierarchical statistical analysis of a training set composed of structures and ER binding data of 313 chemicals. The training dataset comprises inactive, weak and powerful ER binders, and while they represent a variety of chemical classes, these were mainly pesticides and hormone analogues, and were not representative of industrial chemicals or regulated food ingredients. The output is a predicted percentage relative binding affinity to the ER for target substances that fall within the chemical applicability domain of the model. For such limited applicability of the domain, the accuracy of prediction has been reported to be around 84% (Klopman and Chakravarti, 2003). In his paper evaluating the OECD (Q)SAR Application Toolbox for the profiling of ER binding affinities (Level 1 of the OECD CF), Mombelli (2012) notes that the predictive performances of the ER-profiler are lower than that obtainable by utilising tests from Level 2 of the OECD CF, such as the stably

[13] http://ihcp.jrc.ec.europa.eu/our_activities/food-cons-prod/endocrine_disrupters/eas_database

transfected transcriptional activation assay (TA test) for detecting oestrogenic activity (TGs 455, 457). The author concludes however that, even though predictions are not as reliable as the TA test, ER-profiler and more generally (Q)SAR models for binding affinities, can still provide a useful tool for grouping chemicals into categories and/or prioritizing chemicals for experimental testing. They can also play an important role as a first step of an integrated testing strategy for reproductive toxicity. For further discussion, see also the reviews by Benfenati et al. (2005), and Roncaglioni and Benfenati (2008).

The recent conceptual construct of Adverse Outcome Pathways, linking a molecular initiating event with an apical adverse outcome that is relevant to a regulatory decision (Ankley et al., 2010; OECD, 2012e) is being promoted by the OECD as a way to integrate existing knowledge from *in vivo* tests with the results of molecular screening and omics assays, computational predictive methods and other sources of information. Once completed and agreed, they will be included in the OECD (Q)SAR Toolbox, and OECD member countries can propose and initiate work on such AOPs, including endocrine relevant AOPs.

Another (Q)SAR model in the Toolbox is the Danish EPA's Relative ERBA (Q)SAR, which is based on ER binding *in vitro*. The results from this model are based on the (Q)SAR database developed by the Danish EPA to support regulatory assessment of chemicals, and comprises predictions by some 70 models for around 166,000 organic chemicals for a wide range of endpoints.

The Virtual Toxlab also provides a free platform- the OpenVirtual Toxlab (Vedani et al., 2012).

4.2.2.6. Expert systems based on computational models

Expert systems are computational tools that combine the different *in silico* approaches to predict bioactivity of a chemical substance from its structure, and the term is not to be confused with 'expert judgement'. A transparent and data rich regulatory example is that of the US EPA ER expert system. In 2009, the Federal Insecticide, Fungicide, and Rodenticide Act Scientific Advisory Panel (FIFRA SAP) of the US EPA externally reviewed the use of "An Effects Based Expert System to Predict oestrogen Receptor (ER) Binding Affinity for Food Use Inert Ingredients and Antimicrobial Pesticides: Application in a Prioritisation Scheme for Endocrine Disruptor Screening" (U.S. EPA-SAP, 2009a, 2009b). This expert system is designed to support prioritisation of chemicals for screening and testing. A decision based SAR alert system based upon binding to the rainbow trout ER (rtER) was reviewed by the OECD (OECD, 2009b). This expert system is considered to have good concordance with the human and rat ERalpha. The rtER expert system has been incorporated in the OECD (Q)SAR Toolbox, which also contains the ER binding alert developed by OASIS.

The US EPA is targeting over 10,000 chemicals in developing their Endocrine Disruptor Screening Program (EDSP) and the work plan for this project includes *in silico* approaches as a major priority (U.S. EPA-SAP, 2011). Guidelines based upon the OECD validation principles (OECD, 2004) have recently been drawn up for the validation of *in silico* methods for this purpose (U.S. EPA, 2012a, 2012b). The FIFRA SAP is being convened to further address scientific issues associated with prioritizing the EDSP chemicals using computational toxicology tools in early 2013, and the final report is due to be published by May 2013.

Additional tools of relevance include the application of physiologically-based pharmacokinetic (PBPK) modelling in chemical risk assessment (Andersen, 2003; Blaauboer, 2003; Bouvier d'Yvoire et al., 2007; Clewell and Clewell, 2008; Adler et al., 2011), and dose response modelling methods (e.g. Burman and Wiklund, 2011; Plan et al., 2012). However, current research activities do not seem to be optimally directed at developing computational tool alternatives for endocrine-related reproductive and developmental toxicity endpoints in relation to PBPK modelling (Punt et al., 2011). To date both modelling techniques have had far greater application in pharmaceutical research and

clinical development, than for chemical risk assessment, including screening substances for endocrine activity, but remain highly relevant and an important computational area for further development particularly for the integration of data produced from *in vitro/in silico* methods into a biologically meaningful framework and for the extrapolation to *in vivo* conditions (Adler et al., 2011).

4.2.2.7. Summary

With the large diversity of chemicals that might interfere with the endocrine system and the many potential molecular targets, the establishment of a single (Q)SAR for EASs is an unrealistic goal. However, for receptor mediated interactions (e.g. ER, AR, PXR, PPAR, GR, AhR), application of *in silico* methods in sequential/ step-wise approaches (combining relevant and reliable expert systems or (Q)SAR models), can contribute to the WoE basis in the prediction of toxicity to endocrine relevant modalities. *In silico* methods are also useful tools for screening and prioritisation for further testing.

When interrogating such computational tools, a high level of attention needs to be paid to chemistry and biological endpoint data quality and data cleaning considerations, appropriate selection and use of descriptors and statistics. The use of (Q)SAR models, expert systems, category formation tools, as well as the interpretation of the results, require expert knowledge, because each of these tools have their own level of reliability and chemical applicability domain limitations. Additionally, a number of limitations should be kept in mind including the fact that molecules of similar structure sometimes have dissimilar biological activity.

Overall, for molecular initiating event endpoints, such as ER and AR binding and activation, the quality and reliability of the tools are relatively high. However, for *in vivo* tissue, target organ and whole animal endocrine toxicity, such as reproductive and developmental toxicity, these tools have been found to be of limited applicability and low reliability compared to *in vitro* and *in vivo* testing. With respect to future work, it is recommended to explore additional tools, such as the application of PBPK and dose response modelling methods for screening and evaluating substances for endocrine activity.

4.3. General overview of available test methods at CF Levels 2-5, according to endocrine modality and taxa

The purpose of this section is to give a general overview of the availability of test methods from the OECD CF for the identification and characterisation of an EAS/ED. The information is presented according to endocrine modalities that are known to (or could) be sensitive to disruption. Further details on the test methods available for mammals, aquatic organisms and birds are given in sections 4.4, 4.5 and 4.6 respectively. For some endocrine modalities and taxa, no mechanistic and/or apical assays have been developed yet in internationally standardised (validated) test methods. This is probably due to priority setting, as for instance, determined by the relevance for humans and the environment, or for instance the resources and needs in the OECD Member States to do this work. The areas for further development that have been identified are therefore also presented in the following paragraphs.

Table 1 provides a list of selected endocrine modalities/axes/pathways. These are not exhaustive (DRP 178, OECD, 2012b) and it is as yet unknown whether all of them can be affected by endocrine disruption in reality. The table furthermore lists the main taxa on which *in vivo* testing is performed (or may be proposed to be performed), and also includes *in vitro* testing. Information is provided as to whether mechanistic (M) and/or apical (A) tests (either standardised at the international level, or soon

to be so) are available for each of these taxa, and for each of the indicated endocrine modalities/axes/pathways.

Regarding the applicability of test methods across taxa, more is known about mammalian endocrine systems than about equivalent systems in other taxa, although it should be noted that many of these systems have been highly conserved during the evolution of the vertebrates. This means that in some cases, for example with the oestrogen receptor, there is a good species concordance. However, with other receptors, this is not necessarily so and moreover, the downstream effects of hormonal modulation vary considerably between species. The level of available knowledge about the endocrinology of the various taxa (including mammals) is therefore also indicated in Table 1.

4.3.1. Endocrine modalities known to be sensitive to disruption and their corresponding test methods

Most of the knowledge about endocrine disruption has been acquired for substances which interact with oestrogen, androgen, or thyroid hormone systems or affect steroidogenesis in vertebrates (the so-called EATS modalities). Such substances may be able to interact with ligand binding sites (i.e. by mimicking the relevant endogenous agonist (where known)), or block hormone action (i.e. act as an antagonist), or interfere with the endogenous ligand (hormone) synthesis, transport or metabolism. Of these effects in vertebrates, the so-called 'feminising' actions of oestrogens and their mimics are probably the most studied and best understood, having been described in all groups from fish to mammals. Some effects of endocrine disrupting substances on human health have also been described (e.g. diethylstilbestrol and glycyrrhizinic acid). In addition, there is reasonably good knowledge about substances which mimic or antagonise the ecdysteroids and juvenile hormones in insects and crustaceans, because a variety of target-specific insecticides have been explicitly designed to have this type of action in arthropods. Furthermore, it is known that organotin compounds can cause masculinisation in female molluscs, probably by interfering with the retinoid signalling pathway, but possibly also by interfering with steroidogenesis (aromatase inhibition). Finally, vertebrate sex steroids are able to cause adverse effects in molluscs, although the mechanisms of action are also not fully understood.

In vitro EA(T)S tests based on mammalian systems:
The currently (or soon to be) available internationally standardised *in vitro* assays are only applicable for detecting oestrogen/androgen/steroidogenic activity (see Table 1). It should be noted that none of these assays have significant metabolic competence. This means that the metabolic pathways as seen in a whole animal and responsible for bioactivation or de-activation of parent molecules cannot be evaluated with the current *in vitro* assays. However, the OECD is encouraging metabolism components to be added to the available *in vitro* tests (Jacobs et al., 2008; OECD, 2008).

For substances affecting the thyroid hormonal axis, whilst tests providing information on thyroid-related apical endpoints and thyroid-related *in vivo* biomarkers are available, standardised *in vitro* mechanistic screens are still lacking. The OECD is considering this as an area for further work, especially because sufficiently developed tests (e.g. TR binding assay, iodide uptake, thyroid peroxidase inhibition, Thyroid Hormone (TH) transport protein displacement) already exist (see DRP 57, OECD, 2006).

In vitro EA(T)S tests based on other vertebrate systems:
These will be potentially covered in the future when additional test guidelines become available. This will not be soon, since they are not in the current OECD work plan (OECD, 2012d). However, there is

a DRP on Environmental Endocrine Disruptor Screening: the Use of Estrogen and Androgen Receptor Binding and Trans-activation Assays in Fish (see DRP 135, OECD, 2010).

In vivo EATS tests with vertebrates:
With respect to the standardised *in vivo* assays, most of the EATS modalities and their apical effects are detectable using a range of mechanistic and/or apical assays based on mammals, fish and (to a lesser extent) amphibians. There are no standardised assays for these modalities in reptiles, and only apical assays in birds.

Thus, at present, there is still a need to complete validation of several mechanistic and apical *in vivo* tests based on fish, birds and amphibians with sensitivity to EATS modalities, although currently available assays nevertheless allow a fairly effective oestrogen/androgen/steroidogenesis hazard assessment using fish.

In vivo tests with invertebrates:
In invertebrates, no standardised mechanistic assays are available, although apical reproduction assays under development with molluscs have sensitivity to some vertebrate steroids which form part of the EATS modalities, and equivalent assays with insects (and probably crustaceans) are sensitive to ecdysteroids and juvenile hormone mimics. None of these apical tests is able to provide a firm diagnosis of a specific endocrine activity linked to a given adverse effect.

Summary of in vitro and in vivo (non-)validated EA(T)S tests with aquatic organisms and birds:
Considering the standardised assays as a whole, a reasonably complete set of assays is only available for assessing both the mechanistic and apical effects of substances acting by oestrogenic/androgenic/steroidogenic modalities in a single group: fish. In other words, as of now, ecotoxicity testing of EASs is only able to provide a fairly comprehensive assessment of hazards to fish alone, posed by this limited group of substances (although even in fish, there is currently no internationally validated lifecycle test that includes mechanistic endpoints). At present, there is only a validated mechanistic screen (with limited apical endpoints) for thyroid effects in frogs, and it is not possible to measure longer-term apical hazards in amphibians.

Almost the entire focus of aquatic test method development for EASs to date has been on the EATS modalities in vertebrates. Within three years, it is expected that the validated aquatic *in vivo* testing suite for mechanistic and apical testing of EASs acting via the EATS modalities in vertebrate aquatic wildlife will be largely complete for regulatory purposes through the combined use of fish and amphibian assays. A multi-generation test with birds is also expected to be internationally standardised within the same timescale. However, there is little prospect at present of developing a lifecycle test with amphibians.

4.3.2. Non-EATS endocrine modalities not at present known to be sensitive to disruption and for which standardised test methods are not available

Other endocrine modalities have been described in a DPR on The State of the Science on Novel *in vitro* and *in vivo* Screening and Test methods and Endpoints for Evaluating Endocrine Disruptors (DRP 178, OECD, 2012b). As mentioned in the OECD DRP 178, it is possible that a range of additional endocrine modalities in vertebrates (including Hypothalamus-pituitary-adrenocortical axis (HPA) axis; somatotropic axis; retinoid pathway; vitamin D pathway; Peroxisome Proliferator-Activated Receptor (PPAR) signalling pathway; pancreatic signalling; renal signalling) may be susceptible to endocrine disruption (see Table 1). The downstream effects of disruption of these

endocrine pathways are probably detectable in some standardised apical vertebrate assays. For example, evidence for impairment of functional capacity of the endocrine pancreas may be found in vertebrate apical studies (degenerative changes, e.g. concerning the islets of Langerhans may be detected by histopathological examination in the context of OECD TG 408, and changes in blood sugar levels would be covered by clinical chemistry). However, as yet there are no standardised mechanistic assays which are able to identify these modes of action (for the taxa discussed above). However, as discussed in the DRP, endpoints with sensitivity to these modalities could be added to existing *in vivo* vertebrate apical assays, and *in vitro* screens could also be developed and/or standardised.

Another area of lack of knowledge concerns interference with the many peptide hormone pathways in invertebrates. Again, the apical insect or crustacean reproduction and/or lifecycle assays may reveal some of the downstream effects of such interference, but standardised mechanistic assays for these modalities are not available. Furthermore, many other invertebrate phyla (e.g. echinoderms; nemertines) have not been the subject of any standardised test developments, apical or otherwise, although apical reproduction assays for annelids, mites and collembolans have been published by OECD.

The Epigenome and endocrine disruption: Epigenetic modulations underlie critical developmental processes and contribute to determining adult phenotype. Alterations to the phenotype, due to exposure to environmental insults during sensitive periods of development, are mediated through alterations in epigenetic programming in affected tissues. The Annex to DRP 178 (OECD, 2012b) evaluates the potential role of chemical-induced epigenetic modifications to endocrine signalling pathways, during sensitive windows of exposure, as a mechanism of endocrine disruption, and examines the potential methods and assays for assessing such disruption in a screening and testing programme. It concludes that whilst it may be premature to initiate OECD test guideline activity, because of the rapid scientific development in this field, it is important to monitor progress.

In vitro and in vivo (non-)validated non-EATS tests for mammals (human health)

Some of the validated mammalian apical tests may also be sensitive to endocrine modalities other than the EATS modalities (e.g. involving other axes or signalling pathways), but validated *in vitro* and *in vivo* mechanistic screens for such other endocrine modalities have not been developed yet, although work is ongoing (OECD, 2012b). In the short-term, it has been suggested to enhance the available guideline mammalian toxicity studies with the addition of further apical endpoints sensitive to these other modalities (OECD, 2012b).

In vitro and in vivo (non-)validated non-EATS tests for the environment

Some of the validated apical assays using fish may be sensitive to other vertebrate endocrine modalities, although this remains to be demonstrated, and they do not provide mechanistic data. Also the validated arthropod apical assays are expected to be sensitive to ecdysteroid and juvenile hormone disruptors, but again provide no mechanistic information. However, several more assays (described in section 4.5.1) are expected to complete international validation in the next three years to overcome this need for non-EATS mechanistic assays. As with the assays that have already been validated, those still in validation are focused on one or more of the EATS modalities, although it is possible that the apical effects of some other known modalities will be detected by them, depending on whether they are vertebrate- or arthropod-based.

Non-EATS modalities in aquatic vertebrates

Many of these non-EATS modalities have been described in detail (OECD, 2012b), and non-validated mechanistic assays for some already exist, although most are *in vitro* rather than *in vivo*. Therefore,

the major gaps in the aquatic test guideline suite concerns *in vivo* mechanistic assays for the non-EATS modalities in vertebrates. There is a potential need to develop and/or validate a suite of mechanistic *in vitro* and/or *in vivo* assays for aquatic vertebrates, covering non-EATS modalities such as the corticosteroid axis, the somatotropic axis, vitamin D signalling, retinoid signalling and the PPAR pathway. Before such developments take place, however, research is required to decide which (if any) of these types of endocrine modalities can be damaged by chemical exposures of wildlife. If it is concluded that some of these modalities *are* disrupted in the real world, it may become necessary to investigate and possibly augment the responsiveness of some existing *in vivo* assays.

Non-EATS modalities in aquatic invertebrates:

Another major gap in the aquatic test guideline suite is the absence of *in vivo* mechanistic assays for all modalities in invertebrates. To some extent, this development is being held back by the poor knowledge of endocrinology in many invertebrate phyla. Therefore, basic knowledge about the endocrine systems in several invertebrate phyla (*e.g*. molluscs) needs to be improved to permit the development of mechanistic *in vitro* and/or *in vivo* assays based on these taxa. Such mechanistic assay development is probably possible for arthropods at the present state of knowledge with respect to ecdysteroid and juvenile hormone disruptors.

4.3.3. Other general limitations in the available test suite

Sensitive life stages – whole life cycle

In relation to mammals, a limitation of the current suite of test methods available for the identification of EDs (and therefore an area for further developing it) is the lack of a single study involving exposure through the complete life cycle of a mammal, from conception to old age or a single study involving developmental exposure with follow-up into old age.

Limitations of animal models

It has also been suggested that a relevant weakness of current test methods is the limitation of some animal models in relation to certain human endocrine disorders in which EDs have been suggested to play a role, such as some mammary gland tumours and other hormonal cancers, endometriosis, metabolic syndrome and reproductive senescence (Kortenkamp et al., 2011).

The Scientific Committee notes that as the limitations described above are discussed in DRP 178 (OECD, 2012b), their consideration is envisaged at the OECD level.

Despite the fact that the existing internationally standardised assays might miss some endocrine-sensitive endpoints, this should not necessarily lead to the non-identification of EDs. Given the complexity of the endocrine system with its multiple signalling pathways and cross-talks, an ED is expected to produce a pleiotropic response with a range of effects, some of which are likely to be observed in an appropriate guideline study.

Table 1: Degree of knowledge about selected endocrine modalities/axes/pathways and availability (now and in the near future) of internationally standardised mechanistic (**M**) and apical (**A**) methods for identifying and/or characterising the effects of EASs. **A?** = non-validated, but probable apical sensitivity. **n/a** = not applicable. **HPA** = Hypothalamus – Pituitary – Adrenocortical axis; **PPAR** = Peroxisome Proliferator-Activated Receptor signalling pathway. For the endocrine systems covered on the right-hand side of the table as from HPA axis onwards, it is noted that their sensitivity to disruption by realistic exposures is largely unknown at present. Main source of information is the OECD (OECD, 2012a, 2012b, 2012f).

	Knowledge of endocrinology in animal taxa	***Endocrine modalities/axes/pathways and availability of internationally standardised test methods for identifying and characterising EDs***											
		Oestrogen	**Androgen**	**Thyroid**	**Steroidogenesis**	**Ecdysteroids**	**Juvenile hormones**	**HPA axis**	**Somatotropic axis**	**Retinoid pathway**	**Vitamin D pathway**	**PPAR pathway**	**Others (e.g. neuro-peptides in inverts; pancreatic and renal signalling systems in vertebrates)**
***1) In vitro* systems**	n/a	M	M	M?	M								
***2) In vivo* vertebrates**													
Mammals	Good-Excellent	M / A	M / A	M / A	M / A			A?	A?	A?	A?	A?	A?
Birds	Moderate-Good	A	A	A?	A?					A?			A?
Reptiles	Poor-moderate												A?
Amphibians	Moderate-Good	A?	A?	M / A	A?			A?		A?		A?	A?
Fish	Moderate-Good	M / A	M / A	A?	M / A			A?	A?	A?		A?	A?
***3) In vivo* invertebrates**													
Insects	Poor-Moderate					A	A						A?
Crustaceans	Poor-Moderate					A	A						A?
Molluscs	Poor	A	A		A					A			A?
Echinoderms	Poor												
Other phyla	Absent-Poor												

4.4. Available internationally standardised test methods for mammals

The OECD revised CF lists five *in vitro* (mammalian and non-mammalian) mechanistic screens at Level 2, two *in vivo* screens at Level 3, eleven *in vivo* tests at Level 4 and two *in vivo* tests at Level 5. The complete list of the CF is extracted and presented in Appendix C.

The Level 2 *in vitro* mechanistic screens include validated (either by the OECD or US EPA) tests for oestrogen receptor (ER)-mediated, androgen receptor (AR)-mediated and steroidogenesis (S) interference-mediated modalities. These are the ER US EPA OPPTS 890.1250 (U.S. EPA, 2011a) and the AR US EPA OPPTS 890.1150 (U.S. EPA, 2011b) binding assays, the ER transactivation assay (OECD TG 455), the steroidogenesis assay (OECD TG 456) and the aromatase US EPA OPPTS 890.1200 assay (U.S. EPA, 2011c). Level 2 also includes a test for ER agonists and antagonists (OECD TG 457) and assays for thyroid hormone-mediated modalities (TR binding assay, iodide uptake, thyroid peroxidase inhibition, TH transport protein displacement) for which no validated methods are yet available. A validated AR transactivation assay is also not available, but validation work is ongoing.

The Level 3 *in vivo* screening assays include two validated tests, one sensitive to oestrogen agonists/antagonists (Uterotrophic assay in rodents, OECD TG 440) and one sensitive to androgen agonists/antagonists (Hershberger assay in rodents, OECD TG 441).

The Level 4 *in vivo* tests are essentially guideline mammalian toxicity studies, which include apical endpoints sensitive to endocrine disruption. Some of these tests may also include additional biomarkers of endocrine activity (e.g. thyroid and sex steroid hormone levels). The Level 4 tests are: the enhanced 28-day study (OECD TG 407); the 90-day study (OECD TG 408); the one-generation reproduction toxicity study (OECD TG 415); the male pubertal assay (US EPA OPPTS 890.1500, (U.S. EPA, 2011d)); the female pubertal assay (US EPA OPPTS 890.1450, (U.S. EPA, 2011e)); the intact adult male endocrine screening assay (no guideline available); the prenatal developmental toxicity study (OECD TG 414); the chronic toxicity and carcinogenicity studies (OECD TG 451-3); the enhanced reproductive screening test (OECD TG 421); the enhanced combined 28-day/reproductive screening assay (OECD TG 422); and the developmental neurotoxicity study (OECD TG 426). These tests are predominantly responsive to one or more of the EATS modalities, but may be also sensitive to some additional endocrine modalities (e.g. those involving the corticosteroid axis, somatotropic axis, vitamin D signalling, retinoid signalling, pancreatic system, or PPAR pathway – as described in DRP 178 (OECD, 2012b), or other endocrine glands/structures). For example, in the repeated dose toxicity studies (OECD TG 407, 408, 451-3), histopathological investigations of the adrenal gland could provide data on endocrine relevant endpoints involving the corticosteroid axis; or in the reproductive toxicity studies (OECD TG 414, 415, 421, 422, 426), growth evaluation could provide information on apical endpoints involving the somatotropic axis and the retinoid signalling pathway.

The Level 5 *in vivo* tests include two validated reproductive toxicity studies, the extended one-generation study (OECD TG 443) and the two-generation study (OECD TG 416). These tests provide more comprehensive data on endocrine-relevant apical endpoints (predominantly mediated by the EATS modalities but also by other endocrine modalities such as those involving the corticosteroid axis, somatotropic axis, vitamin D signalling, retinoid signalling, PPAR pathway or other endocrine glands/structures) over more extensive parts of the life cycle of the organism, although old age or senescence is generally not covered.

Table 2 lists the *in vivo* mammalian toxicity screens and tests from Levels 3 to 5 of the OECD revised CF, showing their known or potential responsiveness to various endocrine modalities.

Table 2: *In vivo* mammalian toxicity screens and tests listed in the OECD CF, showing their known or potential responsiveness to various (selected and not exhaustive) endocrine modalities/axes/pathways. For each test, its level of the CF is shown: those at Level 3 are suitable for identification of endocrine activity, while those at Levels 4 and 5 are more suitable for hazard identification and characterisation. **M**: screen providing some mechanistic information; **A**: screen or test providing some apical information; **P**: apical endpoints potentially responsive, but not yet fully evaluated.

Endocrine modalities/axes/ pathways	***Assays***														
	Uterotrophic assay (OECD 440) (Level 3)	Hershberger assay (OECD 441) (Level 3)	Enhanced 28-day study (OECD 407) (Level 4)	90-day study (OECD 408) (Level 4)	1-generation study (OECD 415) (Level 4)	Male pubertal assay (US EPA OPPTS 890.1500) (Level 4)	Female pubertal assay (US EPA OPPTS 890.1450) (Level 4)	Intact adult male assay (no TG available) (Level 4)	Prenatal dev tox study (OECD 414) (Level 4)	Chronic tox and carcinogenicity studies (OECD 451-3) (Level 4)	Enhanced reproductive screening assay (OECD 421) (Level 4)	Enhanced combined 28-day/reproductive screening assay (OECD 422) (Level 4)	Developmental neurotoxicity study (OECD 426) (Level 4)	Extended 1-generation study (OECD 443) (Level 5)	2-generation study (OECD 416) (Level 5)
Oestrogen	M		A	A	A		A		A	A	A	A	A	A	A
Anti-estrogens	M		A	A	A		A		A	A	A	A	A	A	A
Androgen	M	M	A	A	A	A		A	A	A	A	A	A	A	A
Anti-androgen		M	A	A	A	A		A	A	A	A	A	A	A	A
Thyroid		M	M/A	A	A	A	A	A	A	A	A	A	A	A	A
Anti-thyroid		M	M/A	A	A	A	A	A	A	A	A	A	A	A	A
Steroidogenesis			A	A	A	A	A	A	A	A	A	A	A	A	A
HPA/Corticosteroid axis			P	P	P				P	P	P	P	P	P	P
Somatotropic axis			P	P	P				P	P	P	P	P	P	P
Vitamin D signalling			P	P	P				P	P	P	P	P	P	P
Retinoid signalling			P	P	P				P	P	P	P	P	P	P
PPAR pathway			P	P	P				P	P	P	P	P	P	P
Other potential endocrine modalities not covered in OECD, 2012b			P	P	P				P	P	P	P	P	P	P

4.5. Available internationally standardised test methods for aquatic organisms

The CF lists five aquatic ecotoxicity tests at Level 3, twelve aquatic tests at Level 4, and seven aquatic tests at Level 5, although several of these are still in the early stages of validation. Table 3 lists the *in vivo* aquatic assays with some sensitivity to EDs; nine have been nationally or internationally validated. Some of the validated assays are essentially mechanistic screens (OECD TG 230; OECD GD 140 (OECD, 2009c)), while others have both mechanistic and apical endpoints (OECD TG 229; OECD TG 231; OECD TG 234), and a third category just have apical endpoints at present (US EPA OPPTS 850.1500 (U.S. EPA, 2004), OECD TG 218/219; OECD TG 211; OECD TG 233).

The assays giving some mechanistic information all use vertebrates (fish or amphibians), and are responsive to one or more of (anti-)oestrogens, (anti-)androgen, (anti-)thyroid, or steroidogenesis-disrupting (EATS) modalities. On the other hand, additional potential modalities (involving *inter alia* the corticosteroid axis; somatotropic axis; vitamin D signalling; retinoid signalling; or PPAR pathway (see DRP 178, OECD, 2012b) may be expected to affect endpoints such as growth, reproduction and development in the US EPA OPPTS 850.1500 fish lifecycle toxicity test (U.S. EPA, 2004). However, this assay would not at present give information about mechanisms, and it has not been validated with these modalities in mind. Furthermore, it is not yet known whether the effects of these modalities occur at realistic levels of exposure, so test developments in this area may be premature. Finally, it is probable that the *Daphnia* reproduction and chironomid lifecycle assays are sensitive to arthropod-specific EDs, but again they do not provide mechanistic data.

Detailed guidance on the interpretation of results from many of the aquatic assays is given in the OECD Fish Toxicity Framework (OECD, 2012g) (see also Appendix C for a short introduction).

The nine aquatic ecotoxicity assays with expected sensitivity to some EDs but with no assigned numbers in Table 3 are still in development and/or validation at OECD. One of these is a mechanistic screen for thyroid-acting substances in frogs, the *Xenopus* Embryo Thyroid Signalling Assay, which uses transgenic *X. laevis* larvae and gives diagnostic results in three days (much more quickly than OECD TG 231). Another amphibian method currently being validated, the Larval Amphibian Growth and Development Assay which exposes *X. laevis* from the larval stage to sexual maturity, is expected to provide mechanistic and apical information about EATS modalities, but does not include reproduction as an endpoint so cannot be considered a full lifecycle test.

There are also two higher-tier fish tests in validation. The first is the Fish Reproduction Partial Lifecycle Test which is essentially an OECD TG 229 followed by TG 234, thus exposing fish (several possible species including *Pimephales promelas* and *Danio rerio*) from reproduction in the F0 generation through to sexual differentiation of the F1 generation. The second is the Medaka Multi-Generation Test (MMGT) which exposes *Oryzias latipes* from reproduction of the F0 fish to reproductive maturity in the F2 generation. Both tests include some mechanistic as well as apical endpoints. One focus of the validation programme is to establish whether the MMGT provides any greater sensitivity to certain EDs than single lifecycle tests. If not, a version of the US EPA OPPTS 850.1500 fish lifecycle toxicity test (U.S. EPA, 2004), with added mechanistic endpoints may be sufficient to provide the high tier information required at Level 5 of the CF.

Several apical invertebrate assays are being validated by OECD. These include Partial and Full Lifecycle Tests with gastropod molluscs (*Potamopyrgus antipodarum* and *Lymnaea stagnalis*), a Copepod Reproduction and Development Test with *Amphiascus tenuiramis*, a Mysid 2-Generation Test with *Americamysis bahia*, and a *Daphnia magna* Multigeneration Test. It is expected that the molluscs will show sensitivity to some vertebrate EATS modalities as well as to others such as disruptors of retinoid signalling (e.g. organotins), while the arthropods are expected to be sensitive *inter alia* to ecdysteroid and juvenile hormone disruptors.

Table 3: Aquatic *in vivo* ecotoxicity screens and tests listed in the OECD CF, showing their known or potential responsiveness to various (selected and not exhaustive) endocrine modalities/axes/pathways. For each test, its Level of the CF is shown: those at Level 3 are suitable for identification of endocrine activity, while those at Levels 4 and 5 are more suitable for hazard identification and characterisation. **M**: screen providing some **mechanistic** information; **A**: screen or test providing some apical information; **P**: apical endpoints potentially responsive, but not yet fully evaluated.

	Vertebrate assays										***Invertebrate assays***							
Endocrine modality/ axes/ pathways	Fish short term reproduction assay (TG 229) (Level 3)	21 d Fish Assay (TG 230) (Level 3)	Androgenised female stickleback screen (GD 140) (Level 3)	Amphibian metamorphosis assay (TG 231) (Level 3)	*Xenopus* embryo thyroid signalling assay (expected to be Level 3)	Fish sexual development test (TG 234) (Level 4)	Larval amphibian growth & development assay (Level 4)	Fish reproduction partial lifecycle test (Level 4)	Fish lifecycle toxicity test (US EPA OPPTS 850.1500) (Level 5)	Medaka multigeneration test (Level 5)	Mollusc partial lifecycle toxicity test (Level 4)	Chironomid toxicity test (TG 218/219) (Level 4)	*Daphnia* reproduction test (TG 211) (Level 4)	Mysid 2-generation toxicity test (Level 5)	Copepod reproduction & development test (Level 5)	Chironomid lifecycle toxicity test (TG 233) (Level 5)	Mollusc lifecycle toxicity test (Level 5)	*Daphnia* multigeneration test (Level 5)
Oestrogen	M/A	M				M/A	M/A	M/A	A	M/A	P						P	
Anti-oestrogens	M/A	M				M/A	M/A	M/A	A	M/A	P						P	
Androgen	M/A	M	M			M/A	M/A	M/A	A	M/A	P						P	
Anti-androgen	M/A	M	M			M/A	M/A	M/A	A	M/A	P						P	
Thyroid				M/A	M		M/A	P	P	P								
Anti-thyroid				M/A	M		M/A	P	P	P								
Steroidogenesis	M/A	M				M/A	M/A	M/A	A	M/A	P						P	
HPA-axis - Corticosteroid axis							P	P	P	P								
Somatotropic axis							P	P	P	P								
Vitamin D							P	P	P	P								
Retinoid signalling							P	P	P	P	P						P	
PPAR pathway							P	P	P	P								
Epigenetic effects										P				P				P
Ecdysteroid system												A	P	A	A	A		A
Juvenile hormone system												A	P	A	A	A		A

4.6. Available internationally standardised test methods for birds

Whilst effects upon predatory birds, particularly eggshell thinning provided early sentinel warnings of the endocrine disrupting effects of some pesticides over 50 years ago, at present there are no finalised, internationally standardised and agreed test methods specifically for assessing the hazards of endocrine disrupting substances to birds. Test guideline work is under development (OECD, 2005b).

The Avian Two Generation Test with the Japanese quail *Coturnix japonica* runs for 21 weeks, from 4 weeks old F0 reproducing adults to 2 weeks old F2 chicks, and hence encompasses more than one complete generation. It is therefore expected to be responsive to most substances with EATS activities. This test is included in the OECD Work plan 2012. Some of the endpoints affected by EATS in this test are also included in an apical test (TG 206).

OECD TG 206 is designed primarily as an apical test for chemicals with suspected reproductive toxicity, but it is not a lifecycle test as it only runs from the stage of pre-laying adults to 14 days old offspring. Furthermore, only the adults are exposed to the test substance (via food), and any effects on sexual development would not be detectable. The endpoints are all apical measures of development, growth or reproduction. Key endpoints which might be affected by EDs include egg production, viability and hatchability.

Detailed information for testing of birds is also given in the DRP on the avian two-generation toxicity test (OECD, 2007b).

Regarding the interpretation of results from the above tests, detailed guidance is given in GD150 (OECD, 2012a) and a table with interpretation results is presented in Appendix C. Briefly, the outcome of the above two tests in combination with other *in vivo* and *in vitro* information provides a tool for assessing whether a compound is a potential endocrine active substance or not.

4.7. Discussion on the appropriateness of available internationally standardised test methods

4.7.1. Considerations specific to endocrine activity and disruption

The appropriateness of test methods regarding the coverage of modalities and taxa is reviewed in the section above (4.3). In conclusion, considerable strides have been made by the OECD in the last decade to develop/standardise *in vitro* and *in vivo* mechanistic and *in vivo* apical assays with sensitivity to EDs. A reasonably complete suite of standardised assays (for endocrine activity and for endocrine hazard identification and/or characterisation) is only available (or will soon be available) for the EATS modalities relevant for mammals and fish, with fewer tests available for birds and amphibians. While downstream effects of disruption of some non-EATS pathways/modalities may be detectable in some standardised apical vertebrate assays, it is important to recognise that standardised mechanistic assays for non-EATS modalities relevant to mammals, fish and other vertebrates are not or not yet available. For invertebrates relevant mechanistic assays are conspicuous by their absence from the OECD testing suite, mainly due to poor understanding of invertebrate endocrinology. Finally, a range of major taxa such as reptiles and echinoderms have not yet been considered by OECD for any endocrine assay development. It is unknown at present whether it will be possible to read-across to untested groups from tests with other taxa.

4.7.2. Considerations not unique to endocrine activity and disruption

Other considerations of appropriateness relate to more general aspects while testing effects of substances (i.e. not exclusively for testing potential endocrine activity/disruption) and are discussed below.

4.7.2.1. Critical windows of susceptibility

This section discusses whether or not available test methods cover appropriately exposure during the critical windows of susceptibility to EASs. It is noted that this issue of critical windows of susceptibility is relevant for all developmental toxicants and not only for EASs.

Hormones, among other factors, are important for the correct development of organs and tissues to take place. Disruption at critical points during such development can result in irreversible changes of the organ/tissue.

In mammals, critical periods of development have been identified at conception, during pregnancy, infancy, childhood and puberty. For a comprehensive review of endocrine mediated effects and the timing of exposure in mammals including humans see Annex I of SAAED (Kortenkamp et al., 2011) and the recent WHO report on possible developmental early effects of endocrine disruptors on child health (WHO, 2012).

Also in other vertebrates such as fish, disturbances at critical periods of development can result in dysfunction and/or disease across the entire lifespan.

In wildlife, there are a number of examples of exposure during critical windows of susceptibility which can lead to irreversible effects. For example, permanent changes in fish sex ratios are induced following early-life stage exposure to some androgen, oestrogen and steroidogenesis disruptors. Thus, exposure of zebrafish embryos and fry to some oestrogens (e.g. 17alpha-ethinylestradiol) can lead to an increase in females in the phenotypic sex ratio when the fish start to differentiate sexually (Andersen et al., 2003). A second example concerns male sticklebacks exposed to an oestrogen as juveniles. Even though their phenotypic sex remains unaltered and despite them being reared to adulthood in clean water, their sexual behaviour is affected and they make fewer nests and rear fewer eggs than normal males (Maunder et al., 2007).

It is widely accepted that, in relation to potential effects from exposure during critical periods of susceptibility, testing *in vivo* is required. This is to encompass sufficiently sensitive endpoints of toxicological relevance during the sensitive life stages that allow judgement of adversity. To avoid the possibility that relevant effects are overlooked, the administration of test compounds needs to address recognised periods of sensitivity and endpoint assessment has to cover all life stages. In the OECD CF for testing and assessment of endocrine disrupting substances, some Level 4 and 5 tests do cover critical periods of development *in utero* and in later life stages. On the other hand, fish lifecycle and multi-generation tests cover all relevant windows of exposure and can be expected to reveal the longer-term effects of even short-term exposures at all stages of the lifecycle.

However, several recent review reports concluded that current mammalian tests do not cover certain endpoints that might be induced by exposure during fetal or pubertal development but emerge later in life like certain cancers (breast, prostate, testis, ovarian and endometrial) and effects on reproductive senescence (EFSA, 2010; Kortenkamp et al., 2011; EEA, 2012; OECD, 2012b). See also section 4.3.3.

4.7.2.2. Combined exposure to multiple substances

In the assessment of combined exposure toxicity involving several substances affecting a common target, the concept of dose addition may be applied to predict the toxicological outcome, assuming that all of the substances to which exposure occurs might contribute to the common effect, depending on their individual potency and their individual concentrations.

The Scientific Committee recognises that exposure to more than one EASs (e.g. from a mixture or from several sources at approximately the same time) could occur in such a way that combined toxicity could arise. EDs can work together to produce additive effects, even when combined at low doses that individually do not produce observable effects, "*making accurate risk assessment difficult or impossible*" for this type of substances (WHO/UNEP, 2013).

The SC recognises that information on interactions at the receptor level may be obtained from *in vitro* studies. However, because of differences in toxicokinetic properties of the various substances, it is not enough to predict the nature of the combined effects from *in vitro* studies, when an intact organism would be exposed to any combination of substances from the same category (with a likely similar mechanism contributing to the specified mode of action).

The SC notes that the current tests are being developed to address single substances. It acknowledges that it is necessary to first develop the tests for single substances with adequate dose response data for reference substances (first for EATS and then for other modalities), before addressing combined exposure to multiple substances.

The SC notes that these considerations are not unique to EASs, but are equally applicable to substances with other mechanisms or modes of action. Considering the complexity of combined exposure to multiple substances, this aspect of hazard and risk assessment of EASs could not be further addressed here. For further information on this issue, see the following existing or soon available publications:

- the opinion of the EC Scientific Committees (SCCS, SCHER, SCENIHR) on Toxicity and Assessment of Chemical Mixtures (SCHER/SCENIHR/SCCS, 2011),
- the EC State of the Art Report on Mixture Toxicity (Kortenkamp et al., 2009),
- the opinion of the EFSA Scientific Panel on Plant Protection products and their Residues (PPR Panel) to evaluate the suitability of existing methodologies and, if appropriate, the identification of new approaches to assess cumulative and synergistic risks from pesticides to human health (EFSA, 2008),
- the opinion of the PPR Panel on risk assessment for a selected group of pesticides from the triazole group to test possible methodologies to assess cumulative effects from exposure through food from these pesticides on human health (EFSA Panel on Plant Protection Products and their Residues, 2009),
- chapter 6 of the EFSA Scientific Opinion on the science behind the development of a risk assessment of Plant Protection Products on bees (*Apis mellifera*, Bombus spp. and solitary bees), which discusses how to take account of cumulative and synergistic effects (EFSA Panel on Plant Protection Products and their Residues, 2012b, and references therein),
- an EFSA Scientific Report reviewing the terminology, methodologies and frameworks developed by national and international agencies for the human risk assessment of combined exposure to multiple chemicals. The publication of this report is expected later in 2013.
- an EFSA PPR Panel Scientific Opinion on the identification of pesticides to be included in cumulative assessment groups on the basis of their toxicological profile. The publication of this opinion is expected later in 2013.

4.7.2.3. Low-dose effects and non-monotonic dose response curves

Conventional *in vivo* toxicity testing in regulatory toxicology and ecotoxicology, e.g. in chronic, sub-chronic or reproductive / developmental studies, involves a high dose and a few lower doses (usually two to four) that cover about a 50- to 100- fold range. The high dose has to be near the maximum tolerated dose, i.e. a dose that causes some sublethal effect in the organism (e.g. loss of body weight in mammals) to reveal the full toxic potential of the substance. With pesticide dossiers, for example, the current practice is normally to select the lowest dose at a level expected not to result in adverse effects. The conventional hazard characterisation approach for threshold effects considers that no detectable adverse effects will occur below this No-Observed-Adverse-Effect-Level (NOAEL) / No-Observed-Effect[14]-Concentration (NOEC) and based on that paradigm, these cut-off levels can be used as Reference Points (also called Points of Departure) for derivation of health-based guidance values in human or ecotoxicological risk assessments.

In the scientific literature on EASs or EDs, an extensive discussion is taking place on two aspects related to the shape of the dose response curve, which could challenge this paradigm:

- The issue of possible effects at low doses (meant as doses or concentrations below those which are considered to be no-effect levels, or levels of exposure which are (far) below the current health-based guidance value (e.g. Tolerable Daily Intake (TDI) or Acceptable Daily Intake (ADI))) has been subject to discussion for more than a decade (Melnick et al., 2002; EFSA, 2010).
- For EASs it has been proposed that monotonic dose response curves cannot be assumed, but rather, that for such substances non-monotonic dose responses should be anticipated. Such non-monotonic dose response curves (NMDRCs) can be characterised by a change in slope direction along the dose interval studied, contrary to conventional dose responses, which show a consistent increase in adversity along the dose range (Vandenberg et al., 2012).

While most reported low dose and NMDRCs findings have been observed with rodents and are therefore connected to human health, studies in molluscs and fish suggest that these findings can also be relevant for environmental organisms (e.g. Jobling et al., 2004).

The term low-dose effects is not synonymous with or equivalent to NMDRC. Many authors use the terms low-dose effects and NMDRCs interchangeably, which creates considerable confusion.

It has been postulated that conventional *in vivo* testing is inadequate for the identification of low-dose effects or NMDRCs displayed by some EDs and for the characterisation of the shape of the dose response curve in the low-dose range, as monotonicity is normally assumed, and doses much lower than the traditional NOAEL/NOEC are usually not investigated (Vandenberg et al., 2012). Since the effects of low doses cannot be predicted by the effects observed at high doses, the authors recommend to use a wider range of doses, extending into the low-dose range, when testing substances for possible endocrine disrupting properties.

In contrast, other authors have contested whether low-dose effects and NMDRCs do exist. This subject has been the topic of intense debate, mainly due to issues relating to reproducibility. Already in 2000, the issue of low-dose effects was addressed in a review by US-NTP (U.S. NTP, 2001; Melnick et al., 2002), in which it was pointed out that these findings could not be replicated and that their

[14] The NOEC in ecotoxicological studies is usually defined by the absence of adverse effects in the same way as the NOAEL in mammalian studies.

toxicological relevance had not been determined (EFSA, 2010). A few examples of reports/reviews casting doubts on the relevance of NMDRCs (Kitchin and Drane, 2005; Mushak, 2007; Chapin et al., 2008; Mayo and Spanos, 2008) were cited in the EFSA report (2010). Thus, it should be considered that assessment of reproducibility of findings that might indicate low-dose effects or NMDRCs, is important to rule out spurious results. In a response to the paper by Vandenberg et al. (2012), Rhomberg and Goodman (2012) argued that Vandenberg et al. have based their arguments on examples which have been questioned by other scientists and do not provide sound evidence to substantiate their views.

During the June 2012 EFSA Scientific Colloquium on low dose responses in toxicology and risk assessment (EFSA, 2012), discussion was conducted on the question of whether there was sufficient scientific evidence for the existence of low-dose effects and NMDRCs. These findings were considered plausible and reliable by some of the participants at the Colloquium, however, no consensus[15] on this was reached amongst the experts participating in the colloquium discussions. It was stated that the quality and robustness of data for studies reporting NMDRCs should be assessed, as for any other studies. During the EFSA Colloquium, it was further suggested to undertake a critical examination of the existing literature to explore the reliability of the low-dose effects and NMDRCs findings, and to assess how they may impact on current hazard and risk assessment approaches and testing strategies (EFSA, 2012).

The US National Institute for Environmental Health Sciences/NIH and the Joint Research Centre's Institute for Health and Consumer Protection organised a low dose effects and non-monotonic dose responses for endocrine active chemicals workshop in Berlin on 11-13 September 2012. This workshop sought to examine the evidence for low dose effects and non-monotonic dose responses in relation to endocrine active chemicals, with the goal of establishing whether the current observations are sufficient to re-examine the ways in which chemicals are tested for endocrine disrupting properties, and how risks to human health may be managed. Most of the participants were in agreement that non-monotonic dose responses do occur and may be expected at some dose ranges for some substances, but the extent to which they might occur at so-called 'low doses' was considered to be a separate issue. There was a suggestion that a definition of 'low dose' would be helpful, as it is currently used with different meanings in different contexts, and that there was also a need to carry out a practical assessment of the type of effects that may be considered adverse, in the context of endocrine disruption.

Most recently, low-dose effects and NMDRCs have been considered in the WHO/UNEP State of the Science of Endocrine Disrupting Chemicals – 2012 report (WHO/UNEP, 2013).

> *"The 'low dose' hypothesis posits that exogenous chemicals that interact with hormone action can do so in a manner that is quite specific such that traditional toxicological endpoints are not sufficient to preclude adverse outcome, and they do so with dose responses that are nonlinear and potentially non-monotonic (Vandenberg et al., 2012*)."

In relation to NMDRCs, a main message of the WHO/UNEP report is that

> "*EDs produce nonlinear dose responses both in vitro and in vivo; these non linear dose responses can be quite complex and often include non-monotonic dose responses. They can be due to a variety of mechanisms; because endogenous hormone levels fluctuate, no threshold can be assumed.*"

Thus on balance, reviewing the recent colloquia, workshop and WHO/UNEP expert report, the debate is evolving in the scientific community as to the existence and/or relevance of low-dose effects and NMDRCs in (eco)toxicology in relation to endocrine disruption or other endpoints/modes of action, but still lacks consensus. More work needs to be conducted to agree the definitions of the respective

[15] It should be noted that the objective of EFSA Scientific Colloquia is to air views in areas for which the scientific thinking is not yet concluded.

terms, and in practical terms to consider whether or how it could impact upon risk assessment (i.e. assessment of dose response relationships for adverse effects) and testing strategies.

Therefore, the Scientific Committee cannot conclude whether the current test methods are adequate to fully define dose response relationships. However, the available information is equally insufficient to conclude that current dose response analysis in regulatory (eco)toxicology should be modified on a routine basis. Nevertheless, on a case-by-case basis, if triggered by unusual findings, an extended dose response analysis could be performed in a second tier.

The SC further notes that, as low-dose effects and NMDRCs are not unique to endocrine activity, these subjects merit a follow-up in a broader context.

4.8. Testing strategies

As mentioned in section 4.1, in principle no single test allows by itself a conclusion that a substance is an ED; all the available information (*in silico*, *in vitro* and *in vivo* data, including observational studies) should be considered. There is therefore a need for further development of testing strategies to generate adequate data for the identification and assessment of endocrine disrupting properties. An example of a test strategy for fish developed at an OECD workshop in 2010, which makes use of some of the tests that are discussed in sections 4.2.2 and 4.5, is described below.

The approach described in the OECD Fish Toxicity Testing Framework (OECD, 2012g) (see Appendix C for details and a decision tree) provides an outline strategy for collecting all the data needed to perform a risk characterisation. It forms a sound basis for developing an EU-wide ecotoxicity testing strategy for substances, which includes full consideration of both endocrine-related and other toxicity. At the present stage of test validation and development, it should however be noted that it will only be possible to design testing strategies sensitive to EDs with EATS modes of action. Schemes sensitive to other possible modes of action cannot be implemented until the ecological importance of such modalities has been evaluated, and suitable *in vitro* and *in vivo* screening assays have been validated. Furthermore, it is apparent that the generic fish testing strategy is data-intensive and might not be applicable to substances for which few data are likely to be available.

5. Elements of hazard characterisation of endocrine disruptors

A hazard-based approach is included in EU regulation on the assessment of certain substances (e.g. active substances in plant protection products), for possible endocrine disruptive properties (see section 1.2). Section 3 discusses the criteria for the identification of a substance as an ED. As hazard identification is only the first step of hazard assessment, with hazard characterisation being the next, the SC examined whether there is a scientific basis to consider critical effect, severity, (ir)reversibility and potency, for the characterisation of EDs.

The REACH Regulation (see Appendix B) provides that substances with endocrine disrupting properties can be identified as "*substances of very high concern (SVHCs)*" and included in the Candidate List of Chemicals , and may subsequently be made subject to the authorisation requirement of the Regulation. Identification of EDs as SVHCs under REACH Article 57(f) requires that the substance has endocrine disrupting properties and that there is scientific evidence of probable serious effects to human health or the environment, which gives rise to an equivalent level of concern to the level of concern of other substances listed in Article 57(a-e) (i.e. Carcinogenic, Mutagenic or toxic for Reproduction (CMR) Category 1A and 1B, Persistent, Bioaccumulative and Toxic (PBT), very Persistent and very Bioaccumulative (vPvB)).

5.1. Critical effect

Hazard identification usually reveals a range of qualitatively different adverse effects of a substance, representing different toxic endpoints such as body weight decrements or organ toxicity, e.g. liver toxicity or neurotoxicity. These adverse effects are elicited in most cases at different dose or concentration levels. During risk assessment, the hazard characterisation step (e.g. establishment of a health-/ecotoxicology-based guidance value) consists of a dose/concentration–response assessment to identify at which dose/concentration adverse effects occur, taking into account uncertainties. The first adverse effect, or its known precursor key event, that occurs in the most sensitive species as the dose rate or concentration of an agent (e.g. a substance) increases is called the 'critical effect'. Dose/concentration-response modelling may be needed to determine which effect results in the lowest reference point, i.e. which is the critical effect. The SC is of the opinion that hazard characterisation should be based on the effect leading to the lowest health/ecotoxicology-based guidance value.

ECETOC (2011) proposed to use the concept of 'critical effect' in identifying a chemical as an ED for regulatory purposes with the rationale that, if the endocrine-mediated adverse effects occur within a range up to 10 times higher than the critical effect, the substance is then considered as an ED. However, the SC disagrees with the idea that a substance can be identified as an ED only when the endocrine-mediated adverse effects occur within a certain range of the critical effect. According to the agreed definition and criteria for EDs, all substances with the ability to cause adverse effects consequent to an endocrine mode of action are to be regarded as EDs, independently from critical effect considerations. The proposal by ECETOC goes beyond the hazard identification of EDs and falls at the interface between science, policy and risk management, and hence outside the remit of EFSA.

5.2. Severity / (ir)reversibility / potency

Severity and (ir)reversibility
Severity may describe the magnitude of an adverse effect and/or the qualitative nature of the effect, and thus may be associated with the typical manifestations of certain endpoints (e.g. developmental effects or cancer).

(Ir)reversibility considerations may contribute to judgment of severity, although an effect may be regarded as being severe without being irreversible. Reversibility implies that recovery of the individual or population may occur after cessation of exposure to the toxic substance. The SC considers that to inform whether exposure to a substance represents a toxicological risk[16], severity and (ir)reversibility should be evaluated in relation to degree and timing of exposure.

Potency
Potency in general terms can be seen as a descriptor for the relationship between the biological or biochemical effect elicited by a chemical substance and the dose or concentration necessary to achieve this effect. Potency is therefore a measure of a substance's activity or strength to produce the effect, and is part of the dose response considerations in the hazard characterisation of a substance as described above. Potency is usually considered in a context, e.g. in comparison of substances

[16] The problem formulation discussion between the risk assessor and the risk manager aims at defining the parameters within which risk is assessed; this includes among other the identification of protection goals and the definition of adverse effects. The risk assessor will then be able to characterise the toxicological risk of a given substance and its level of concern in relation to the pre-agreed protection goals.

displaying the same effect. Relative potency of an ED which for example mimics a hormone should also be considered in comparison to that of the natural endogenous hormones.

Potency values established in single *in vitro* assays may differ from the dose-effect relationships observed in the intact animal. In accordance with the criteria for defining an ED, which link alteration of functions of the endocrine system to the adverse effects in the intact organism, potency in the present opinion therefore refers to potency associated with *in vivo* adverse effects.

Potency for a particular endpoint *in vivo* may depend not only on the degree of exposure (the dose), but also on the duration and timing of exposure. Thus, for the establishment of potency values for EDs, critical periods of development (studies covering different life stages) and the duration of exposure should be taken into account. The SC is of the opinion that, to assess whether or not a (predefined) level of concern is reached for an ED, potency should not be used alone but should take account of actual or predicted exposure.

It is the opinion of the SC that, if regulation of identified EDs is to be based on a level of concern, whether or not this level of concern is reached, can only be determined by risk assessment. This should take actual or predicted exposure into account, and consider the whole body of evidence in a combined manner to characterise the risk.

Whether hazard characterisation criteria alone, or risk assessment should be used for defining the level of concern for identified EDs for further regulatory measures is beyond the scope of this opinion and is a risk management decision.

CONCLUSIONS

Definition of endocrine disruptor and adverse effect

- The Scientific Committee (SC) defined an endocrine active substance (EAS) as:

  "*any chemical that can interact directly or indirectly with the endocrine system, and subsequently result in an effect on the endocrine system, target organs and tissues.*" (EFSA, 2010).

- It should be pointed out that the range of substances that may be identified as EASs is contingent on our increasing understanding of the endocrine system.

- The SC concludes that the WHO/IPCS 2002 definition of endocrine disruptor (ED) and the WHO/IPCS 2009 definition of adverse effect should be adopted as a basis for the criteria for the identification of EDs:

  - *"An endocrine disrupter is an exogenous substance or mixture that alters function(s) of the endocrine system and consequently causes adverse health effects in an intact organism, or its progeny, or (sub)populations."* (WHO/IPCS, 2002).

  - *"An adverse effects is a change in the morphology, physiology, growth, development, reproduction, or, life span of an organism, system, or (sub)population that results in an impairment of functional capacity, an impairment of the capacity to compensate for additional stress, or an increase in susceptibility to other influences."* (WHO/IPCS, 2009).

Criteria for identifying endocrine disruptors

- An ED is defined by three criteria i) the presence of an adverse effect in an intact organism or a (sub)population; ii) the presence of an endocrine activity; and iii) a plausible causal relationship between the endocrine activity and the adverse effect.

    - Assessment of adversity is not unique to endocrine-related adverse effects. Scientific criteria for assessment of adversity have not been generally defined. In general, but not always, transient, inconsistent and minor fluctuations at the biochemical and molecular level may be considered adaptive, i.e. non-adverse. Changes at the cell-, organ-, organism-, or (sub) population level resulting in pathology or functional impairment *in vivo*, as well as altered timing of development, may be considered adverse. It is therefore difficult to propose ED-specific criteria for adversity and expert judgement is therefore required to assess on a case-by-case basis the (eco)toxicological relevance of such changes and when the biological threshold between endocrine modulation and adverse effect has been crossed.

    - Endocrine-related effects observed secondary to marked toxicity caused by a non-endocrine mode of action, should not be considered specific, genuine endocrine disrupting effects.

    - Endocrine activity information could be obtained from existing information, read-across, *in silico* tools, *in vitro* and *in vivo* screening assays (Levels 1, 2 and 3) of the current OECD Conceptual Framework (CF) or from other mechanistic investigations. These mechanistic studies should be evaluated on their merits on a weight-of-evidence basis.

    - A prerequisite for an EAS to be regarded as an ED is the need to identify the adverse effect. For this, test methods with apical endpoints (Levels 4 and 5 of the CF) can be used together with existing information, read-across and other *in vivo* (eco)toxicity tests that provide information on apical endpoints.

    - There must be a reasonable evidence base for a biologically plausible causal relationship between the induced endocrine activity and the adverse effect(s) seen in an intact organism, or its progeny, or (sub)population. Evidence for this relationship should be obtained from the OECD CF or from other investigations and assessed on a weight-of-evidence (WoE) basis.

Availability and appropriateness of test methods for identifying and characterising effects mediated by endocrine active substances

Test methods for both mammalian and non-mammalian endocrine modalities, most of which have been developed and subsequently (internationally) standardised (and validated) in the framework of OECD programmes, were considered with respect to appropriateness. Methods used predominantly for academic research purposes were not specifically considered.

- *Computational toxicology and non-test methods:* Application of *in silico* methods in sequential/step wise approaches, combining relevant and reliable expert systems and/or (Quantitative) Structure Activity Relationship models, can contribute on a WoE basis to the prediction of endocrine activity. Overall, for molecular initiating event endpoints, such as

oestrogen receptor (ER) and androgen receptor (AR) binding and activation, the quality and reliability of the tools are relatively high. However, for the prediction of endocrine related toxicity *in vivo*, such as reproductive and developmental toxicity, these tools have been found to be of limited applicability and low reliability.

- *In vitro oestrogen, androgen, thyroid or steroidogenesis (EATS) tests:* The currently (or soon to be) available internationally standardised *in vitro* assays based on mammalian systems are only applicable for detecting oestrogen, androgen, or steroidogenic activity. For substances affecting the thyroid hormonal axis, whilst tests providing information on thyroid-related apical endpoints and thyroid-related *in vivo* biomarkers are available, *in vitro* mechanistic screens relevant to the thyroid are still lacking. Work is ongoing at OECD level on thyroid. *In vitro* EA(T)S tests based on other vertebrate systems will be potentially covered in the future when additional test guidelines become available.

- *In vivo EATS tests for vertebrates:* With respect to the standardised mechanistic and/or apical *in vivo* assays, most of the EATS modalities and their expected apical effects are only detectable in mammalian and fish (and to a lesser extent amphibian) assays. There are no standardised assays for these modalities for reptiles, and only apical assays for birds.

- *Non-EATS tests for vertebrates:* Some of the standardised mammalian apical tests may also be sensitive to other than EATS modalities (e.g. involving other axes or signalling pathways), but internationally standardised *in vitro* and *in vivo* mechanistic screens for these modalities have not been developed yet, although work is ongoing. Regarding the environment, some of the fish apical assays standardised for oestrogen, androgen or steroidogenic modalities may be sensitive to other vertebrate endocrine modalities.

- *EATS and non-EATS tests for invertebrates:* Standardised arthropod apical assays are sensitive to ecdysteroid and juvenile hormone disruptors, but provide no mechanistic information. There are no standardised mechanistic assays for any modalities in invertebrates, although apical reproduction assays under development with molluscs have sensitivity to some vertebrate steroids. None of these apical tests is able to provide a firm diagnosis of a specific endocrine activity linked to a given adverse effect.

In principle, no single assay is likely to provide all the information needed to decide whether a substance is an ED because of the need to provide both mechanistic information showing how the substance interacts with the endocrine system, and apical information describing the adverse effects this interaction may cause.

Appropriateness of available internationally standardised test methods

Taken together, but bearing in mind the recommendations made below, a reasonably complete suite of standardised assays (for testing the effects of EASs) is available (or will soon be available) for the EATS modalities relevant for mammals and fish, with fewer tests available for birds and amphibians. While downstream effects of disruption of some non-EATS pathways/modalities may be detectable in some standardised apical vertebrate assays, it is important to recognise that standardised mechanistic assays for non-EATS modalities relevant to mammals, fish and other vertebrates are not yet available. For invertebrates relevant mechanistic assays are conspicuous by their absence from the OECD testing suite, mainly due to poor understanding of invertebrate endocrinology. Finally, a range of major taxa, e.g. reptiles or echinoderms have not yet been considered by the OECD for any endocrine assay development. It is unknown at present whether it will be possible to read-across to untested groups from tests with other taxa.

The SC discussed briefly the following aspects related to the testing of substances. The Committee notes that these considerations are not unique to EASs, but are equally applicable to substances with other mechanisms of action.

- *Critical windows of susceptibility:* To assess potential effects from exposure during critical windows of susceptibility, testing *in vivo* is required to encompass sensitive life stages. The SC noted that some tests in the OECD CF do cover exposure during critical periods of development *in utero*. However, current mammalian tests may not cover effects that might be induced by exposure during fetal or pubertal development, but may emerge during later life stages. Fish lifecycle tests cover all relevant windows of exposure and can be expected to reveal the longer-term effects of developmental exposures at all stages of the lifecycle.

- *Combined exposure to multiple substances:* The SC recognises that exposure to multiple EASs may occur in such a way that combined toxicity could arise. The issue of mixture toxicity resulting from combined exposure to multiple substances will be addressed by EFSA in a separate activity.

- *Low-dose effects and non-monotonic dose response curves (NMDRCs):* The SC notes the lack of consensus in the scientific community as to the existence and/or relevance of low-dose effects and NMDRCs in (eco)toxicology in relation to endocrine disruption, or other endpoints/modes of actions.

Elements for hazard characterisation of endocrine disruptors

The SC examined in section 5 whether there is a scientific basis to include aspects considered during hazard characterisation, such as critical effect, severity, (ir)reversibility and potency in defining the level of concern for identified EDs.

- *Critical effect:* The SC is of the opinion that hazard characterisation (e.g. establishment of a health/ecotoxicology-based guidance value) should be based on the effect leading to the lowest health/ecotoxicology-based guidance value, irrespective of the mode of action. Such a health/ecotoxicology-based guidance value would also protect against endocrine-mediated effects occurring at higher doses.

- *Severity, (ir)reversibility and potency:* The SC considers that to inform on a level of concern for EASs, severity, (ir)reversibility and potency should be evaluated in relation to degree, timing and duration of exposure. Levels of concern are not determined exclusively by risk assessment but also by protection goals set by the risk management.

In conclusion, endocrine disruptors, of natural or synthetic origin, can be identified according to three criteria: endocrine activity, adversity of effects and a plausible link between endocrine activity and adverse effect. The SC considers that a reasonably complete suite of assays is (or will soon be) available to identify and characterise the important hazards of EATS substances in mammals and fish, with fewer tests available for birds and amphibians. Furthermore, these evaluation methods should, in principle, be fit for the purpose of establishing safe doses/concentrations of EDs if (1) certain aspects (e.g. follow up of exposure at critical windows of susceptibility to later life stages) are addressed and (2) used with all available information in a WoE approach. It should also be noted that standardised mechanistic assays for non-EATS modalities relevant to mammals, fish and other vertebrates are not yet available. For invertebrates, relevant mechanistic assays are lacking from the OECD testing suite. Finally, a range of major taxa e.g. reptiles or echinoderms, have not yet been considered by OECD for any endocrine assay development.

Furthermore, to inform on risk and level of concern for the purpose of risk management decisions it is the opinion of the SC that risk assessment (taking into account hazard and exposure data/predictions) makes best use of available information. EDs can therefore be treated like most other substances of concern for human health and the environment, i.e. be subject to risk assessment and not only to hazard assessment.

RECOMMENDATIONS

Recommendations specific to endocrine active substances

- The following areas were identified as requiring further development regarding test methods:
    - Additional tools, such as the application of physiologically based pharmacokinetic and dose response modelling should be further considered as possible methods to contribute to the screening / evaluation of substances for endocrine activity under Level 1 of the OECD CF.
    - Whilst the suite of standardised assays for testing the effects of EASs is reasonably comprehensive for the EATS modalities relevant for mammals and fish, there are fewer tests available for birds and amphibians or other taxa.
    - In relation to mammals, there is a lack of a single study involving exposure through the complete life cycle, from conception to old age or a single study involving developmental exposure with follow-up into old age.
    - Improvement of predictive models in relation to certain human endocrine disorders in which EDs have been suggested to play a role.

- While downstream effects of disruption of some non-EATS pathways/modalities may be detectable in some standardised apical vertebrate assays, mechanistic assays for non-EATS modalities relevant to mammals, fish and other vertebrates are not or not yet available. Given the costs involved for the development of new standardised test methods, further information on the relevance of possible endocrine disruption of non-EATS modalities in real life (humans and the environment) is needed to prioritise future test development. Further research is therefore needed on whether exposures to substances, which could affect non-EATS modalities, are associated with adverse effects in humans or in the environment.

- The SC also underlined the need for further development of testing strategies to generate adequate data for the identification and assessment of endocrine disrupting properties. An example has been developed in outline for fish species by the OECD.

General recommendations, not unique to endocrine active substances

Although mentioned in the context of EDs, the SC recommends as a follow up activity to clarify the following issues in broader context and in further detail:

- Biological thresholds and criteria for adversity vs. physiological modulation / homeostatic responses,

- Combined exposure to multiple substances,

- Non-monotonic dose response curves.

REFERENCES

Adler S, Basketter D, Creton S, Pelkonen O, van Benthem J, Zuang V, Andersen KE, Angers-Loustau A, Aptula A, Bal-Price A, Benfenati E, Bernauer U, Bessems J, Bois FY, Boobis A, Brandon E, Bremer S, Broschard T, Casati S, Coecke S, Corvi R, Cronin M, Daston G, Dekant W, Felter S, Grignard E, Gundert-Remy U, Heinonen T, Kimber I, Kleinjans J, Komulainen H, Kreiling R, Kreysa J, Leite SB, Loizou G, Maxwell G, Mazzatorta P, Munn S, Pfuhler S, Phrakonkham P, Piersma A, Poth A, Prieto P, Repetto G, Rogiers V, Schoeters G, Schwarz M, Serafimova R, Tahti H, Testai E, van Delft J, van Loveren H, Vinken M, Worth A and Zaldivar JM, 2011. Alternative (non-animal) methods for cosmetics testing: current status and future prospects-2010. Archives of Toxicology, 85, 367-485.

Akahori Y, Nakai M, Yakabe Y, Takatsuki M, Mizutani M, Matsuo M and Shimohigashi Y, 2005. Two-step models to predict binding affinity of chemicals to the human estrogen receptor alpha by three-dimensional quantitative structure-activity relationships (3D-QSARs) using receptor-ligand docking simulation. SAR and QSAR in Environmental Research, 16, 323-337.

Allen Y, Scott AP, Matthiessen P, Haworth S, Thain JE and Feist S, 1999. Survey of estrogenic activity in United Kingdom estuarine and coastal waters and its effects on gonadal development of the flounder *Platichthys flesus*. Environmental Toxicology and Chemistry, 18, 1791-1800.

Andersen L, Holbech H, Gessbo A, Norrgren L and Petersen G, 2003. Effects of exposure to 17 alpha-ethinylestradiol during early development on sexual differentiation and induction of vitellogenin in zebrafish (*Danio rerio*). Comparative Biochemistry and Physiology, 134, 365-374.

Andersen ME, 2003. Toxicokinetic modeling and its applications in chemical risk assessment. Toxicology Letters, 138, 9-27.

Ankley GT, Bennett RS, Erickson RJ, Hoff DJ, Hornung MW, Johnson RD, Mount DR, Nichols JW, Russom CL, Schmieder PK, Serrrano JA, Tietge JE and Villeneuve DL, 2010. Adverse outcome pathways: a conceptual framework to support ecotoxicology research and risk assessment. Environmental Toxicology and Chemistry, 29, 730-741.

Bars R, Broeckaert F, Fegert I, Gross M, Hallmark N, Kedwards T, Lewis D, O'Hagan S, Panter GH, Weltje L, Weyers A, Wheeler JR and Galay-Burgos M, 2011. Science based guidance for the assessment of endocrine disrupting properties of chemicals. Regulatory Toxicology and Pharmacology, 59, 37-46.

Benfenati E, Roncaglioni A, Boriani E, Porcelli C, Spreafico M and Lo Piparo E, 2005. Validation of selected, non-commercial (Q)SAR models for Estrogen Receptor and Androgen Receptor binding. Final report of JRC contract CCR.IHCP.C430414.X0. 96 pp. Available from: http://ihcp.jrc.ec.europa.eu/our_labs/predictive_toxicology/information-sources/qsar-document-area/Final_report_Mario_Negri.pdf.

Bern HA, Mills KT, Hatch DL, Ostrander PL and Iguchi T, 1992. Altered mammary responsiveness to estradiol and progesterone in mice exposed neonatally to diethylstilbestrol. Cancer Letters, 63, 117-124.

Blaauboer BJ, 2003. Biokinetic and toxicodynamic modelling and its role in toxicological research and risk assessment. Alternatives to laboratory animals : ATLA, 31, 277-281.

Boobis AR, Cohen SM, Dellarco V, McGregor D, Meek ME, Vickers C, Willcocks D and Farland W, 2006. IPCS framework for analyzing the relevance of a cancer mode of action for humans. Critical Reviews in Toxicology, 36, 781-792.

Boobis AR, Doe JE, Heinrich-Hirsch B, Meek M, Munn S, Ruchirawat M, Schlatter J, Seed J and Vickers C, 2008. IPCS framework for analyzing the relevance of a noncancer mode of action for humans. Critical Reviews in Toxicology, 38, 87-96.

Bouvier d'Yvoire M, Prieto P, Blaauboer BJ, Bois FY, Boobis A, Brochot C, Coecke S, Freidig A, Gundert-Remy U, Hartung T, Jacobs MN, Lave T, Leahy DE, Lennernas H, Loizou GD, Meek B,

Pease C, Rowland M, Spendiff M, Yang J and Zeilmaker M, 2007. Physiologically-based Kinetic Modelling (PBK Modelling): meeting the 3Rs agenda. The report and recommendations of ECVAM Workshop 63. Alternatives to Laboratory Animals, 35, 661-671.

Bradford Hill A, 1965. The Environment and Disease: Association or Causation? Proceedings of the Royal Society of Medicine, 58, 295-300.

Burman CF and Wiklund SJ, 2011. Modelling and simulation in the pharmaceutical industry - some reflections. Pharmaceutical Statistics, 10, 508-516.

Castello P and Worth A, 2011. Information sources and databases on Endocrine Active Substances. JRC Technical Notes CT 31854. 28 pp. Available from: http://ihcp.jrc.ec.europa.eu/our_activities/food-cons-prod/endocrine_disrupters/eas_database/info-sources-databases-endocrine-active-substances.

Chapin RE, Adams J, Boekelheide K, Gray LE, Hayward SW, Lees PS, McIntyre BS, Portier KM, Schnorr TM, Selevan SG, Vandenbergh JG and Woskie SR, 2008. NTP-CERHR expert panel report on the reproductive and developmental toxicity of bisphenol A. Birth defects research. Part B, Developmental and Reproductive Toxicology, 83, 157-395.

CHEM Trust 2011a. CHEM Trust's Contribution to the Ongoing Debate on Criteria for EDCs. 38 pp. Available from: http://www.chemtrust.org.uk/documents/CHEM%20Trust%20Position%20on%20EDC%20Criteria%20-%20Sept11.pdf.

CHEM Trust 2011b. A CHEM Trust and HEAL Briefing: Challenges and solutions in the regulation of chemicals with endocrine disrupting properties. 8 pp. Available from: http://www.chemtrust.org.uk/documents/Criteria%20Briefing%20CT&HEAL%20FINAL.pdf.

Clewell RA and Clewell HJ, 3rd, 2008. Development and specification of physiologically based pharmacokinetic models for use in risk assessment. Regulatory toxicology and pharmacology : RTP, 50, 129-143.

Colborn T and Corlie C, 1992. Advances in Modern Environmental Toxicology, Volume XXI, Chemically-induced alterations in sexual and functional development: The wildlife/human connection. Princeton Scientific Publishing Co. Inc., New Jersey, USA, 403 pp.

D'Ursi P, Salvi E, Fossa P, Milanesi L and Rovida E, 2005. Modelling the interaction of steroid receptors with endocrine disrupting chemicals. BMC Bioinformatics, 6 Suppl 4, S10.

Dybdahl M, Nikolov NG, Wedebye EB, Jonsdottir SO and Niemela JR, 2012. QSAR model for human pregnane X receptor (PXR) binding: screening of environmental chemicals and correlations with genotoxicity, endocrine disruption and teratogenicity. Toxicology and Applied Pharmacology, 262, 301-309.

Dybing E, Doe J, Groten J, Kleiner J, O'Brien J, Renwick AG, Schlatter J, Steinberg P, Tritscher A, Walker R and Younes M, 2002. Hazard characterisation of chemicals in food and diet. dose response, mechanisms and extrapolation issues. Food and Chemical Toxicology, 40, 237-282.

EC (European Commission), 1997. Proceedings of the European Workshop on the Impact of Endocrine Disrupters on Human Health and Wildlife. Weybridge, UK. 58 pp. Available from http://ec.europa.eu/environment/chemicals/pdf/workshop_report.pdf.

EC (European Commission), 1999. Community Strategy for Endocrine Disrupters - a range of substances suspected of interfering with the hormone systems of humans and wildlife. Available from: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:1999:0706:FIN:EN:PDF.

EC (European Commission), 2000. First report on the harmonisation of risk assessment procedures - Part 1 : The Report of the Scientific Steering Committee's Working Group on Harmonisation of Risk Assessment Procedures in the Scientific Committees advising the European Commission in the area of human and environmental health. Part 2 : Appendices. 26-27 October 2000. Scientific Steering Committee, Health and Consumer Safety Directorate. Available from: http://ec.europa.eu/food/fs/sc/ssc/out82_en.html.

ECETOC (European Centre for Ecotoxicology and Toxicology of Chemicals), 2009. Guidance on Identifying Endocrine Disrupting Effects. Technical Report No 106, 133 pp.

ECETOC (European Centre for Ecotoxicology and Toxicology of Chemicals), 2011. Risk Assessment of Endocrine Disrupting Chemicals. Workshop Report No 21, 32 pp.

EEA (European Environment Agency), 2012. The Impacts of Endocrine Disrupters on Wildlife, People and their Environments – The Weybridge+15 (1996–2011) Report. Technical Report No 2/2012. 116 pp.

EFSA (European Food Safety Authority), 2008. Opinion of the Scientific Panel on Plant Protection products and their Residues to evaluate the suitability of existing methodologies and, if appropriate, the identification of new approaches to assess cumulative and synergistic risks from pesticides to human health with a view to set MRLs for those pesticides in the frame of Regulation (EC) 396/2005. The EFSA Journal, 704, 1-84.

EFSA (European Food Safety Authority), 2009a. Guidance of the Scientific Committee on Transparency in the Scientific Aspects of Risk Assessments carried out by EFSA. Part 2: General Principles. The EFSA Journal, 1051, 1-22.

EFSA (European Food Safety Authority), 2009b. Existing approaches incorporating replacement, reduction and refinement of animal testing: applicability in food and feed risk assessment. The EFSA Journal, 1052, 1-77.

EFSA (European Food Safety Authority), 2010. Scientific Report of the Endocrine Active Substances Task Force. EFSA Journal, 8(11):1932, 59 pp.

EFSA (European Food Safety Authority), 2012. EFSA Scientific Colloquium XVII: Low-dose-response in toxicology and risk assessment. 64 pp. Available from: http://www.efsa.europa.eu/en/supporting/doc/353e.pdf.

EFSA Panel on Plant Protection Products and their Residues (PPR), 2009. Scientific Opinion on Risk Assessment for a Selected Group of Pesticides from the Triazole Group to Test Possible Methodologies to Assess Cumulative Effects from Exposure through Food from these Pesticides on Human Health. EFSA Journal, 7(9):1167, 104 pp.

EFSA Panel on Plant Protection Products and their Residues (PPR), 2012a. Scientific Opinion on Evaluation of the Toxicological Relevance of Pesticide Metabolites for Dietary Risk Assessment. EFSA Journal, 10(7):2799, 187 pp.

EFSA Panel on Plant Protection Products and their Residues (PPR), 2012b. Scientific Opinion on the science behind the development of a risk assessment of Plant Protection Products on bees. (Apis mellifera, Bombus spp. and solitary bees). EFSA Journal, 10(5):2668, 275 pp.

EFSA Scientific Committee, 2011. Statistical Significance and Biological Relevance. EFSA Journal, 9(9):2372, 17 pp.

Foster PM, 2005. Mode of action: impaired fetal leydig cell function - effects on male reproductive development produced by certain phthalate esters. Critical Reviews in Toxicology, 35, 713-719.

Gehring U, Casas M, Brunekreef B, Bergstrom A, Bonde JP, Botton J, Chevrier C, Cordier S, Heinrich J, Hohmann C, Keil T, Sunyer J, Tischer CG, Toft G, Wickman M, Vrijheid M and Nieuwenhuijsen M, 2013. Environmental exposure assessment in European birth cohorts: results from the ENRIECO project. Environmental Health, 12, 8.

Hartung T, Bremer S, Casati S, Coecke S, Corvi R, Fortaner S, Gribaldo L, Halder M, Hoffmann S, Roi AJ, Prieto P, Sabbioni E, Scott L, Worth A and Zuang V, 2004. A modular approach to the ECVAM principles on test validity. Alternatives to Laboratory Animals, 32, 467-472.

Jacobs MN, 2004. *In silico* tools to aid risk assessment of endocrine disrupting chemicals. Toxicology, 205, 43-53.

Jacobs MN, Dickins M and Lewis DF, 2003. Homology modelling of the nuclear receptors: human oestrogen receptorbeta (hERβ), the human pregnane-X-receptor (PXR), the Ah receptor (AhR) and the constitutive androstane receptor (CAR) ligand binding domains from the human oestrogen receptor α (hERα) crystal structure, and the human peroxisome proliferator activated receptor alpha (PPARα) ligand binding domain from the human PPARγ crystal structure. The Journal of Steroid Biochemistry and Molecular Biology, 84, 117-132.

Jacobs MN, Janssens W, Bernauer U, Brandon E, Coecke S, Combes R, Edwards P, Freidig A, Freyberger A, Kolanczyk R, Mc Ardle C, Mekenyan O, Schmieder P, Schrader T, Takeyoshi M and van der Burg B, 2008. The use of metabolising systems for *in vitro* testing of endocrine disruptors. Current Drug Metabolism, 9, 796-826.

Jameson JL, 2010. Harrisson's Endocrinology, 2nd edition. Mc Graw-Hill, New York, USA, 560 pp.

Japanese Ministry of the Environment 2005. MOE's Perspectives on endocrine disrupting effects of substances-ExTEND 2005. 36 pp. Available from: http://www.env.go.jp/en/chemi/ed/extend2005_full.pdf.

Jensen GE, Niemela JR, Wedebye EB and Nikolov NG, 2008. QSAR models for reproductive toxicity and endocrine disruption in regulatory use-a preliminary investigation. SAR and QSAR in Environmental Research, 19, 631-641.

Jobling S, Casey D, Rogers-Gray T, Oehlmann J, Schulte-Oehlmann U, Pawlowski S, Baunbeck T, Turner AP and Tyler CR, 2004. Comparative responses of molluscs and fish to environmental estrogens and an estrogenic effluent. Aquatic Toxicology, 66, 207-222.

JRC (Joint Research Center), 2011. Applicability of (Q)SAR analysis in the evaluation of developmental and neurotoxicity effects for the assessment of the toxicological relevance of metabolites and degradates of pesticide active substances for dietary risk assessment. External Scientific Report. EFSA (European Food Safety Authority). Available from: www.efsa.europa.eu/en/supporting/pub/169e.htm.

Kavlock RJ, Daston GP, DeRosa C, Fenner-Crisp P, Gray LE, Kaattari S, Lucier G, Luster M, Mac MJ, Maczka C, Miller R, Moore J, Rolland R, Scott G, Sheehan DM, Sinks T and Tilson HA, 1996. Research needs for the risk assessment of health and environmental effects of endocrine disruptors: a report of the U.S. EPA-sponsored workshop. Environmental Health Perspectives, 104 Suppl 4, 715-740.

Kidd KA, Blanchfield PJ, Mills KH, Palace VP, Evans RE, Lazorchak JM and Flick RW, 2007. Collapse of a fish population after exposure to a synthetic estrogen. Proceedings of the National Academy of Sciences of the United States of America, 104, 8897-8901.

Kitchin KT and Drane JW, 2005. A critique of the use of hormesis in risk assessment. Human & Experimental Toxicology, 24, 249-253.

Klopman G and Chakravarti SK, 2003. Structure–activity relationship study of a diverse set of estrogen receptor ligands (I) using MultiCASE expert system. Chemosphere, 51, 445-459.

Kortenkamp A, Backhaus T and Faust M, 2009. State of the Art Report on Mixture Toxicity. Final Report of EU Project Contract 070307/2007/485103/ETU/D.1. 391 pp. Available from: http://ec.europa.eu/environment/chemicals/pdf/report_Mixture%20toxicity.pdf.

Kortenkamp A, Martin O, Faust M, Evans R, McKinlay R, Orton F and Rosivatz E, 2011. State-of-the-Art Assessment of Endocrine Disrupters. Final Report of EU Project Contract 070307/2009/550687/SER/D3. 135 pp. Available from: http://ec.europa.eu/environment/endocrine/documents/4_SOTA%20EDC%20Final%20Report%20V3%206%20Feb%2012.pdf.

Lo Piparo E and Worth A, 2010. Review of QSAR Models and Software Tools for Predicting Developmental and Reproductive Toxicity. JRC Scientific and Technical Reports EUR 24522 EN. 32 pp.

Main KM, Skakkebaek NE, Virtanen HE and Toppari J, 2010. Genital anomalies in boys and the environment. Best practice & research. Clinical endocrinology & metabolism, 24, 279-289.

Martina CA, Weiss B and Swan SH, 2012. Lifestyle behaviors associated with exposures to endocrine disruptors. Neurotoxicology, 33, 1427-1433.

Maslankiewicz L, Hulzebos EM, Vermeire TG, Muller JJA and Piersma AH, 2005. Can chemical structure predict reproductive toxicity? RIVM Report No. 601200005, 76 pp.

Matthiessen P and Gibbs PE, 1998. Critical appraisal of the evidence for tributyltin mediated endocrine disruption in mollusks. Environmental Toxicology and Chemistry, 17, 37-43.

Maunder RJ, Matthiessen P, Sumpter JP and Pottinger TG, 2007. Impaired reproduction in three-spined sticklebacks exposed to ethinyl estradiol as juveniles. Biology of Reproduction, 77, 999-1006.

Mayo D and Spanos A, 2008. Risks to health and risks to science: the need for a responsible "bioevidential" scrutiny. Human and Experimental Toxicology, 27, 621-625.

Melnick R, Lucier G, Wolfe M, Hall R, Stancel G, Prins G, Gallo M, Reuhl K, Ho SM, Brown T, Moore J, Leakey J, Haseman J and Kohn M, 2002. Summary of the National Toxicology Program's report of the endocrine disruptors low-dose peer review. Environmental Health Perspectives, 110, 427-431.

Mombelli E, 2012. Evaluation of the OECD (Q)SAR Application Toolbox for the profiling of estrogen receptor binding affinities. SAR and QSAR in Environmental Research, 23, 37-57.

Mushak P, 2007. Hormesis and its place in nonmonotonic dose-response relationships: some scientific reality checks. Environmental Health Perspectives, 115, 500-506.

Newbold RR, Padilla-Banks E, Jefferson WN and Heindel JJ, 2008. Effects of endocrine disruptors on obesity. International Journal of Andrology, 31, 201-208.

NRC (National Research Council), 2007. Toxicity Testing in the 21st Century: A Vision and a Strategy. The National Academies Press, Washington, D.C., USA.

OECD (Organisation for Economic Co-operation and Development), 1992. OECD Guideline for Testing of Chemicals: Test No 210: Fish, Early-Life Stage Toxicity Test.

OECD (Organisation for Economic Co-operation and Development), 1998. OECD Guideline for Testing of Chemicals: Test No 409: Repeated Dose 90-Day Oral Toxicity Study in Non-Rodents.

OECD (Organisation for Economic Co-operation and Development), 2004. OECD Principles for the Validation, for Regulatory Purposes, of (Quantitative) Structure-Activity Relationship Models. 2 pp. Available from: http://www.oecd.org/env/ehs/risk-assessment/37849783.pdf.

OECD (Organisation for Economic Co-operation and Development), 2005a. Series on Testing and Assessment: No 34: Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment. ENV/JM/MONO(2005)14, 96 pp.

OECD (Organisation for Economic Co-operation and Development), 2005b. OECD Guideline for Testing of Chemicals: proposal for a new test guideline, Avian Two-generation Toxicity Test in the Japanese Quail. Draft, Available from: http://www.epa.gov/endo/pubs/edmvac/2gen_guide_gd_draft1.pdf.

OECD (Organisation for Economic Co-operation and Development), 2006. Series on Testing and Assessment: No 57: Detailed Review Paper: Thyroid Hormone Disruption Assays.

OECD (Organisation for Economic Co-operation and Development), 2007a. Series on Testing and Assessment: No 80: Guidance on Grouping of Chemicals. ENV/JM/MONO(2007)28, 99 pp.

OECD (Organisation for Economic Co-operation and Development), 2007b. OECD Guideline for Testing of Chemicals: Test No 74: Detailed Review Paper for Avian Two-generation Toxicity Testing. ENV/JM/MONO(2007)21, 167 pp.

OECD (Organisation for Economic Co-operation and Development), 2008. OECD Guideline for Testing of Chemicals: Test No 97: Detailed Review Paper on the Use of Metabolising Systems for In Vitro Testing of Endocrine Disruptors. ENV/JM/MONO(2008)24, 95 pp.

OECD (Organisation for Economic Co-operation and Development), 2009a. Series on Testing and Assessment: No 102: The Guidance Document for Using the OECD (Q)SAR Application Toolbox to Develop Chemical Categories According to the OECD Guidance on Grouping Chemicals. ENV/JM/MONO(2009)5, 118 pp.

OECD (Organisation for Economic Co-operation and Development), 2009b. Series on Testing and Assessment: No 111: Report of the Expert Consultation to Evaluate an Estrogen Receptor Binding Affinity Model for Hazard Identification. ENV/JM/MONO(2009)33, 119 pp.

OECD (Organisation for Economic Co-operation and Development), 2009c. OECD Guidance Document No 140: The 21-Day Androgenised Female Stickleback Endocrine Screening Assay.

OECD (Organisation for Economic Co-operation and Development), 2010. Series on Testing and Assessment: No 135: Detailed Review Paper on Environmental Endocrine Disruptor Screening: The Use of Estrogen and Androgen Receptor Binding and Transactivation Assays in Fish. ENV/JM/MONO(2010)34, 64 pp.

OECD (Organisation for Economic Co-operation and Development), 2012a. Series on Testing and Assessment: No 150: Guidance Document on Standardised Test Guidelines for Evaluating Chemicals for Endocrine Disruption. ENV/JM/MONO(2012)22, 524 pp.

OECD (Organisation for Economic Co-operation and Development), 2012b. Series on Testing and Assessment: No 178: Detailed Review Paper on the State of the Science on Novel *In vitro* and *In vivo* Screening and Testing Methods and Endpoints for Evaluating Endocrine Disruptors. ENV/JM/MONO(2012)23, 213 pp.

OECD (Organisation for Economic Co-operation and Development), 2012c. Series on Testing and Assessment: No 181: Guidance Document on Standardised Test Guidelines for Evaluating Chemicals for Endocrine Disruption: Case Studies Using Example Chemicals. ENV/JM/MONO(2012)34, 316 pp.

OECD (Organisation for Economic Co-operation and Development), 2012d. Work Plan for the Test Guidelines Programme (TGP). 24 pp. Available from: http://www.oecd.org/env/chemicalsafetyandbiosafety/testingofchemicals/August%20Work%20plan%20for%20the%20Test%20guidelines%20programme%20June%202013.pdf.

OECD (Organisation for Economic Co-operation and Development), 2012e. Draft Proposal for a template and guidance on developing and assessing the completeness of Adverse Outcome Pathways. 17 pp. Available from: http://www.oecd.org/chemicalsafety/testing/49963554.pdf.

OECD (Organisation for Economic Co-operation and Development), 2012f. Information on OECD Work Related to Endocrine Disrupters. 8 pp. Available from: http://www.oecd.org/chemicalsafety/testingofchemicals/50067203.pdf.

OECD (Organisation for Economic Co-operation and Development), 2012g. Series on Testing and Assessment: No 171: Fish Toxicity Testing Framework. ENV/JM/MONO(2012)16, 174 pp.

PAN Europe, 2011. PAN Europe position paper on criteria for endocrine disrupting pesticides. 8 pp. Available from: http://www.pan-europe.info/News/PR/110525.html.

Plan EL, Maloney A, Mentre F, Karlsson MO and Bertrand J, 2012. Performance comparison of various maximum likelihood nonlinear mixed-effects estimation methods for dose-response models. The AAPS Journal, 14, 420-432.

Punt A, Schiffelers MJ, Jean Horbach G, van de Sandt JJ, Groothuis GM, Rietjens IM and Blaauboer BJ, 2011. Evaluation of research activities and research needs to increase the impact and applicability of alternative testing strategies in risk assessment practice. Regulatory toxicology and pharmacology : RTP, 61, 105-114.

Rhomberg LR and Goodman JE, 2012. Low-dose effects and nonmonotonic dose-responses of endocrine disrupting chemicals: has the case been made? Regulatory Toxicology and Pharmacology, 64, 130-133.

Roncaglioni A and Benfenati E, 2008. *In silico*-aided prediction of biological properties of chemicals: oestrogen receptor-mediated effects. Chemical Society Reviews, 37, 441-450.

SCENIHR (Scientific Committee on Emerging and Newly Identified Health Risks), 2012. Memorandum on the use of the scientific literature for human health risk assessment purposes – weighing of evidence and expression of uncertainty. 46 pp. Available from: http://ec.europa.eu/health/scientific_committees/emerging/docs/scenihr_s_001.pdf.

SCHER/SCENIHR/SCCS (Scientific Committee on Health and Environmental Risks/Scientific Committee on Emerging and Newly Identified Health Risks/Scientific Committee on Consumer Safety), 2011. Toxicity and Assessment of Chemical Mixtures. 50 pp. Available from: http://ec.europa.eu/health/scientific_committees/environmental_risks/docs/scher_o_155.pdf.

Swan SH, Main KM, Liu F, Stewart SL, Kruse RL, Calafat AM, Mao CS, Redmon JB, Ternand CL and Sullivan S, 2005. Decrease in anogenital distance among male infants with prenatal phthalate exposure. Environmental Health Perspectives, 113, 1056-1061.

Toppari J, 2008. Environmental endocrine disrupters. Sexual Development, 2, 260-267.

U.S. EPA-SAP (U.S. Environmental Protection Agency-FIFRA Scientific Advisory Panel), 2009a. An Effects-based Expert System to Predict Estrogen Receptor Binding Affinity for Food Use Inert Ingredients and Antimicrobial Pesticides: Application in a Prioritization Scheme for Endocrine Disruptor Screening, Meeting Minutes. EPA-HQ-OPP-2009-0322, 70 pp.

U.S. EPA-SAP (U.S. Environmental Protection Agency-FIFRA Scientific Advisory Panel), 2009b. An Effects-based Expert System to Predict Estrogen Receptor Binding Affinity for Food Use Inert Ingredients and Antimicrobial Pesticides: Application in a Prioritization Scheme for Endocrine Disruptor Screening, Meeting Materials. Available from: http://www.regulations.gov/#!docketDetail;D=EPA-HQ-OPP-2009-0322.

U.S. EPA-SAP (U.S. Environmental Protection Agency-FIFRA Scientific Advisory Panel), 2011. Integrated Approaches to Testing and Assessment Strategies: Use of New Computational and Molecular Tools. FIFRA SAP. Final FIFRA SAP IATA Meeting Report May 2011. EPA-HQ-OPP-2011-0284-0054, 81 pp.

U.S. EPA (U.S. Environmental Protection Agency), 2004. Ecological Effects Test Guidelines OPPTS 850.1500 Fish Life Cycle Toxicity.

U.S. EPA (U.S. Environmental Protection Agency), 2011a. Standard Evaluation Procedure OCSPP 890.1250: Estrogen Receptor Binding Assay Using Rat Uterine Cytosol (ER-RUC). 12 pp.

U.S. EPA (U.S. Environmental Protection Agency), 2011b. Standard Evaluation Procedure OCSPP 890.1150: Androgen Receptor Binding Assay (Rat Ventral Prostate Cytosol). 12 pp.

U.S. EPA (U.S. Environmental Protection Agency), 2011c. Standard Evaluation Procedure OCSPP 890.1200: Aromatase Assay (Human Recombinant). 11 pp.

U.S. EPA (U.S. Environmental Protection Agency), 2011d. Standard Evaluation Procedure OCSPP 890.1500: Pubertal Development and Thyroid Function in Intact Juvenile/Peripubertal Male Rats Assay. 19 pp.

U.S. EPA (U.S. Environmental Protection Agency), 2011e. Standard Evaluation Procedure OCSPP 890.1450: Pubertal Development and Thyroid Function in Intact Juvenile/Peripubertal Female Rats Assay. 19 pp.

U.S. EPA (U.S. Environmental Protection Agency), 2012a. Endocrine Disruptor Screening Program: Universe of Chemicals and General Validation Principles. November 2012. 17 pp. Available from: http://www.epa.gov/endo/pubs/edsp_chemical_universe_and_general_validations_white_paper_11_12.pdf.

U.S. EPA (U.S. Environmental Protection Agency), 2012b. Endocrine Disruptor Screening Program: Universe of Chemicals. November 2012. 177 pp. Available from: http://www.epa.gov/endo/pubs/edsp_chemical_universe_list_11_12.pdf.

U.S. FDA (U.S. Food and Drug Administration), online. Endocrine Disruptor Knowledge Base (EDKB). Available from: http://www.fda.gov/scienceresearch/bioinformaticstools/endocrinedisruptorknowledgebase/default.htm.

U.S. NTP (U.S. National Toxicology Programme), 2001. National Toxicology Program's Report of the Endocrine Disruptors Low-Dose Peer Review. Available from http://ntp-server.niehs.nih.gov/htdocs/liason/LowDoseWebPage.html.

Van Drie JH, 2003. Pharmacophore Discovery. In: Computational Medicinal Chemistry for Drug Discovery. Eds Bultnick P, De Winter H, Langenaeker W, Tollenare JP. CRC Press, New York, USA, 1196 pp.

Vandenberg LN, Colborn T, Hayes TB, Heindel JJ, Jacobs DR, Jr., Lee DH, Shioda T, Soto AM, vom Saal FS, Welshons WV, Zoeller RT and Myers JP, 2012. Hormones and endocrine-disrupting chemicals: low-dose effects and nonmonotonic dose responses. Endocrine Reviews, 33, 378-455.

Vedani A, Dobler M and Smiesko M, 2012. VirtualToxLab - a platform for estimating the toxic potential of drugs, chemicals and natural products. Toxicology and Applied Pharmacology, 261, 142-153.

Vinggaard AM, Niemela J, Wedebye EB and Jensen GE, 2008. Screening of 397 chemicals and development of a quantitative structure-activity relationship model for androgen receptor antagonism. Chemical Research in Toxicology, 21, 813-823.

Vrijheid M, Casas M, Bergstrom A, Carmichael A, Cordier S, Eggesbo M, Eller E, Fantini MP, Fernandez MF, Fernandez-Somoano A, Gehring U, Grazuleviciene R, Hohmann C, Karvonen AM, Keil T, Kogevinas M, Koppen G, Kramer U, Kuehni CE, Magnus P, Majewska R, Andersen AM, Patelarou E, Petersen MS, Pierik FH, Polanska K, Porta D, Richiardi L, Santos AC, Slama R, Sram RJ, Thijs C, Tischer C, Toft G, Trnovec T, Vandentorren S, Vrijkotte TG, Wilhelm M, Wright J and Nieuwenhuijsen M, 2012. European birth cohorts for environmental health research. Environmental Health Perspectives, 120, 29-37.

WHO (World Health Organization), 2012. Endocrine disrupters and child health: Possible developmental early effects of endocrine disrupters on child health. 84 pp. Available from: http://www.who.int/ceh/publications/endocrine_disrupters_child/en/index.html.

WHO/IPCS (World Health Organization/International Programme on Chemical Safety), 2002. Global Assessment of the State-of-the-science of Endocrine Disruptors. WHO/PCS/EDC/02.2, 180 pp.

WHO/IPCS (World Health Organization/International Programme on Chemical Safety), 2009. Principles and Methods for the Risk Assessment of Chemicals in Food. Environmental Health Criteria 240. 689 pp. Available from: http://www.who.int/foodsafety/chem/principles/en/index1.html.

WHO/UNEP (World Health Organization/United Nations Environment Programme), 2013. State of the Science of Endocrine Disrupting Chemicals - 2012. 296 pp. Available from: http://www.who.int/mediacentre/news/releases/2013/hormone_disrupting_20130219/en/index.html

Worth A, Fuart-Gatnik M, Lapenna S and Serafimova R, 2011. Applicability of QSAR analysis in the evaluation of developmental and neurotoxicity effects for the assessment of the toxicological relevance of metabolites and degradates of pesticide active substances for dietary risk assessment. EFSA External Scientific Report. EFSA (European Food Safety Authority).174 pp. Available from: http://www.efsa.europa.eu/de/supporting/doc/169e.pdf.

Zoeller RT, Brown TR, Doan LL, Gore AC, Skakkebaek NE, Soto AM, Woodruff TJ and Vom Saal FS, 2012. Endocrine-disrupting chemicals and public health protection: a statement of principles from The Endocrine Society. Endocrinology, 153, 4097-4110.

APPENDICES

A. OVERVIEW OF CURRENT STATE, RECENT REPORTS AND ACTIVITIES

The EFSA Technical Report on Endocrine Active Substances (EASs) (EFSA, 2010) provides an overview of European and international activities on EASs until 2010. This section aims at updating the reader on developments that occurred between 2010 and 2013.

A.1 European Commission (EC)[17]

In March 1999, the Scientific Committee for Toxicity, Ecotoxicity and the Environment issued a report on 'Human and Wildlife Health Effects of Endocrine Disrupting Chemicals, with Emphasis on Wildlife and on Ecotoxicology Test Methods' (EC, 1999). The report identified a 'potential global problem' for wildlife.

Against this background, a 'Community Strategy for Endocrine Disrupters' was adopted by the EC in December 1999 (EC, 1999)[18]. The objectives were to identify the problem of endocrine disruption, its causes and consequences and to identify appropriate policy action on the basis of the precautionary principle in order to respond quickly and effectively to the problem, thereby alleviating public concern. Four key elements were identified:

- the need for further research,
- the need for international co-ordination,
- the need for communication to the public,
- the need for policy action.

On this basis a set of appropriate short, medium and long term actions was recommended, among which the establishment of a priority list of substances for further evaluation of their role in endocrine disruption and the development of international cooperation on this issue.

Need for further research

Since the start of the 4th Framework Programme (FP) in 1994, through FP5, FP6 and the ongoing FP7[19], the EC funds research on endocrine disruption. Up to now more than 80 projects were launched focusing on endocrine disruptors (EDs) identification, risk assessment, education and information on chemicals as contaminants in the food chain[20]. Annex 1 – section 9.1 of the State of the Art of the Assessment of Endocrine Disruptors (SAAED) (Kortenkamp et al., 2011) provides a recent review of EU projects with potential relevance to the following criteria:

- Assay development / validation
- New insights in mechanisms of EDs
- Comparison and correlation of assays
- Identification of new chemicals with endocrine activity
- Data relevant to exposure assessment
- Data on the occurrence of conditions in humans
- Field studies of effects in wildlife
- Cross cutting issues

[17] The text for section A1 has been taken from relevant EC webpages
[18] See http://ec.europa.eu/environment/endocrine/strategy/index_en.htm.
[19] http://cordis.europa.eu/fp7/home_en.html
[20] http://ec.europa.eu/research/endocrine/index_en.html

Need for international coordination

The EC participated in the OECD working group of the National Co-ordinators for the Test Guidelines Programme (WNT) and also in its subsidiary body the Endocrine Disruptors Testing and Assessment Advisory Group (EDTA AG). This activity resulted in the publication of the OECD guidance document on standardised test guidelines for evaluating chemicals for endocrine disruption (OECD, 2012a).

Some ten years ago the EC provided financial contribution to the publication of a comprehensive report on *Global Assessment of the State-of-the-Science of Endocrine Disruptors*, which was published in 2002 by the IPCS (WHO/IPCS, 2002). Recently, the United Nations Environment Programme (WHO/UNEP) and WHO have published a 10-year update of this report and the EC has been involved (WHO/UNEP, 2013).

Need for policy action

In 2009, a study on the State of the Art of the Assessment of Endocrine Disruptors has been commissioned by the Directorate-General for the Environment to provide a basis for:

i) the development of scientific criteria for the identification of EDs, and
ii) the review and possible revision of the Community Strategy on EDs.

The study was finalised by the contractor at the end of January 2012 and the resulting report (Kortenkamp et al., 2011) published on DG Environment's website.

The report summarises advances in the state of the science since 2002 and maps out ways of dealing with endocrine disruptors in important pieces of EU chemicals regulation, such as the Plant Protection Product Regulation (EC) No 1107/2009, the Biocides Regulation (EC) No 528/2012 and the Chemicals Regulation (EC) No 1907/2006. The report comes to a number of conclusions and recommendations for the assessment of EDs:

- The definition for endocrine disrupting chemicals developed by WHO/IPCS is generally accepted as being applicable to both human health and ecotoxicological hazard and risk assessment.
- Internationally agreed and validated test methods (OECD) for the identification of EDs are generally regarded as useful, but it is acknowledged that they capture only a limited range of the known spectrum of endocrine disrupting effects. Considerable gaps exist for the identification of chemicals that can affect wildlife taxa.
- For a wide range of endocrine disrupting effects, agreed and validated test methods do not exist. This introduces considerable uncertainties, with the likelihood of overlooking harmful effects in humans and wildlife. Until better tests become available, hazard and risk identification has to rely also on epidemiological approaches.
- The information and testing requirements laid down in important pieces of EU chemicals regulation do not capture the range of endocrine disrupting effects that can be measured with internationally agreed and validated test methods. Testing with the most sensitive and appropriate methods currently available and with exposure regimens that cover periods of heightened susceptibility during critical life stages is not conducted.
- An overview of proposals for regulating EDs by EU Member States and other organisations revealed some commonalities and areas of agreement. However, controversy remains regarding the proposals to deal with EDs on the basis of potency-based cut-off values derived from the CLP Regulation (EC) No 1272/2008.
- Defining EDs for regulatory purposes will have to rely on criteria for adversity and endocrine-related modes of action. A decision tree approach is developed that proceeds in a step-wise manner by excluding substances that neither produce adverse effects, nor show endocrine-related modes of action. In the absence of appropriate evidence, relevance should be assumed by default.

- The final regulatory decision rests on a consideration of the toxicological profile of the substances in a weight-of-evidence (WoE) approach that still has to be developed. This WoE approach will have to consider potency together with other factors such as severity and specificity of effect and irreversibility. Rigid potency-based cut-off values as decisive decision criteria are not recommended.
- Procedures that incentivise the provision of data in the case of data gaps are suggested. There are still enormous knowledge gaps that need to be addressed through research and development projects.

On 11 and 12 June 2012, the EC organised a conference on "Endocrine Disruptors: Current challenges in science and policy"; the presentations and discussions covered the effects of EDs on human health and the environment, the risks, the identification of EDs and policy objectives.

The outcome of the conference as well as the above-mentioned Kortenkamp et al. report will feed into the review of the European Commission's current strategy on endocrine disruptors. It will also provide input to the Commission's upcoming proposal for criteria for the identification of substances with endocrine disrupting properties (planned for end of 2013).
To fulfil these tasks, the EC DG Environment established an ad-hoc advisory group of Member States and Commission Services and Agencies (autumn 2010) and an expert sub-group on endocrine disruptors (November 2011). The ad-hoc advisory group is used for:

- information exchange on EDs
- bringing science on EDs and chemical's policy together
- discussing horizontal aspects of regulation on EDs
- providing orientation to the Commission on development and implementation of EU policy in this field

The Expert sub-group is consulted to:

- exchange information on detailed technical and scientific issues on EDs
- support the ad-hoc group in technical and scientific issues

The subgroup will produce a report providing definition and clarifications of basic terms related to EDs, proposing factors for the identification, characterisation and categorisation of EDs, discussing WoE considerations to categorise a substance as an ED, and suggesting a scheme for the evaluation of EDs for regulatory purposes. The report was not yet finalised at the time this opinion was adopted.

A.2 European Environment Agency (EEA)

In the 1996 Weybridge meeting on EDs (EC, 1997), the problem of EDs was first comprehensively discussed by both European and United States regulatory authorities. Much focus was placed on oestrogenic compounds, and especially on receptor-mediated effects.

Since then, substantial EU funds (i.e. over EUR 150 million spent until 2011 have been allocated to research into endocrine disruptors and their effects, and the WHO and the OECD have addressed the problem in many ways. Scientific progress over the last decade has expanded the scope considerably: it includes EDs that affect other hormone systems, e.g. the thyroid; EDs with new modes of action, e.g. inhibitors of endogenous hormone production or metabolism; and target tissues for EDs other than those in the reproductive system, such as the brain and cardiovascular system. The Weybridge+10 workshop organised by the Academy of Finland, the EC DG Research and EEA in 2006 aimed to evaluate the impacts of this extensive research and to determine future goals in the areas of human and wildlife health effects, mechanisms of biological actions and models, exposures, risks, and policy options.

The Weybridge+15 (1996-2011) report (EEA, 2012) reviewed key conclusions, challenges and recommendations that have been drawn from the research over the last 15 years (and discussed in the plenary session at the Weybridge+10 meeting). It concludes, among other, that endocrine disruption is a real phenomenon likely affecting both human and wildlife populations globally, but a much better

understanding of the role of chemicals as causal factors of a wider range of endocrine diseases and disorders is needed. Screening tests and technologies (e.g. 'omics') to detect/predict endocrine disruptive mode of action exist but there are still inadequacies in some areas. More knowledge on what types of EDs are most likely to be affecting humans and wildlife (which are most prevalent and potent?) is needed. Considerable evidence suggests that low-dose effects of EDs often cannot be predicted from high-dose testing. These 'low-dose effects' of EDs have come under intense scrutiny as they oppose traditional toxicology paradigms. The different possible meanings of the term 'low dose' can cause confusion. Non-monotonic dose–response curves for EDs, including 'U-shaped' or 'inverted-U-shaped' curves have been described, thus calling into question the appropriateness of assuming monotonicity as a basis for chemical risk assessments of EDs.

A.3 European Parliament (EP)

The EP Policy Department on Economic and Scientific Policy dealing with Environment, Public Health and Food Safety organised a workshop on 'Endocrine Disruptors and Impact on Health' in Brussels on 18 September 2012. The workshop was organised in the context of the self-initiative taken by the ENVI Committee to better understand the impacts of EDs on health and to provide input into the ongoing policy discussions at EU-level. The outcome of the workshop, as well as the above-mentioned previous work made on the ED issue by the EC and other EU bodies were then considered to prepare a draft report that includes a motion for an EP Resolution on the protection of public health from endocrine disruptors[21]. The latter calls on the Commission to submit as soon as possible proposals for comprehensive criteria together with testing and information requirements for chemicals on the commercial market, and for EU legislation to make clear what is regarded as a substance with endocrine-disrupting properties; it advocates considering the introduction of 'endocrine disruptor' as a regulatory hazard class, and calls on the Commission to ensure that the criteria for identifying EDs are applied horizontally to all current and future legislation, and that appropriate testing requirements for the identification of substances with endocrine-disrupting properties are introduced in all relevant EU legislation. This draft report was adopted by the ENVI Committee on 23 January 2013 with a series of amendments (more than 150 amendments were tabled). The ENVI report notably calls for fast measures to protect vulnerable groups such as children, young people and pregnant women.
A final debate on the basis of the ENVI report took place in plenary session of the European Parliament on 12 March 2013. The report was adopted on 14 March 2013.

A.4 European Food Safety Authority (EFSA)

Scientific report of the Endocrine Active Substances Task Force

An EFSA technical report developed by a cross-EFSA task force was published on 30 November 2010 to clarify the state-of-play on EASs and to make recommendations for scientific and communication issues (EFSA, 2010). Discussions within the Scientific Committee and the EFSA Advisory Forum had called for the development of a common approach within EFSA towards EASs. Both specific issues and new regulations made it necessary to follow up on developments with the EU bodies, Member States, and internationally, in order to avoid diverging assessment approaches and the duplication of work. The proposed actions for EFSA are to contribute to the work in progress under the auspices of the EC. The development of a generally accepted risk assessment methodology was identified as a challenge due to the complexity of the issues involved. Here, the task force recommends that EFSA continues its activities aimed at developing harmonised methodologies for risk assessment of combined exposures to EASs in food. EFSA should continue to build a dialogue to develop a common strategy with the EC, other EU bodies, Member States' Competent Authorities, international organisations and partners, as well as external experts and stakeholders on the before mentioned issues. EFSA should also work with the experts in its Advisory Group on Risk Communications in conjunction with the communication experts from Member States, and continues to monitor and

[21] Available at: http://www.europarl.europa.eu/meetdocs/2009_2014/documents/envi/pr/912/912390/912390en.pdf

analyse media and stakeholder developments, in order to define a strategy for communications addressing both the collective group and specific EASs.

Colloquium on low dose responses

EFSA held its 17th Scientific Colloquium on low dose response in toxicology and risk assessment on 14 June 2012 in Parma, where Scientists debated low-dose hypothesis.

Over two days, 100 scientific international experts exchanged views and debated the possible health effects of low levels of certain chemicals (the 'low-dose hypothesis') and the current and future challenges these pose for food and feed risk assessment. The Colloquium attracted risk assessors, risk managers, scientists and stakeholders from 21 countries, including 12 EU Member States, 4 candidate countries, Japan, Norway, Russia, Switzerland and the United States. Prominent toxicologists, endocrinologists and biochemists from academia, industry and public health authorities took part, including representatives of several European National Competent Authorities, the European Commission, the Joint Research Centre (JRC), the Scientific Committee on Emerging and Newly Identified Health Risks (SCENIHR), the European Chemicals Agency (ECHA) and the U.S. Food and Drug Administration (U.S. FDA).

The Colloquium came at a potentially critical juncture in a scientific debate that has been gaining prominence since the 1990s: increasing numbers of studies address effects of chemical substances at low doses, mainly those substances referred to as EASs or EDs. According to the low-dose hypothesis, these substances may cause adverse effects at low doses but not necessarily at all higher doses. They do not therefore follow the classical (or 'monotonic') dose response curve, showing a greater likelihood of an adverse effect at higher doses. Rather they may show a different kind of dose response curve, e.g. a U-shaped curve with responses both at low- and high-dose levels but not in intermediate ranges. Such a dose response curve is termed a non-monotonic dose response curve (NMDRC). Such findings challenge current concepts in chemical risk assessment.

As yet no scientific consensus has been reached as to the validity of the low-dose hypothesis. Recently, it has been claimed that a large number of new studies provide further support for this hypothesis. The relevance of these findings and the way they could impact on assessment of the possible risks of chemicals in food and feed were the main topic of this event.

Following an introductory session where speakers summarised the current debate, participants were divided into four discussion groups each focusing on a specific key issue: the nature of an effect and the assessment of adversity; dose response relationships; the evidence for NMDRCs; the challenges for risk assessment. Related aspects were covered during the course of the debates, including effects of EASs, testing methods and strategies, and modelling techniques for predicting biological responses.

The objective of EFSA's scientific colloquia is to bring together international experts from different sectors for an open scientific debate on key issues; they are organised so as to provide ample opportunity for the exchange of views.

Presentations given and summary report of the colloquium are available at EFSA's website[22] .

A.5 US NIEHS/NIH and EC Joint Research Centre IHCP

In Berlin on 11-13 September 2012, the US National Institute for Environmental Health Sciences/NIH and the Joint Research Centre's Institute for Health and Consumer Protection organised a workshop on low dose effects and non-monotonic dose responses for endocrine active chemicals: science to practice

[22] http://www.efsa.europa.eu/en/events/event/120614.htm

workshop. This workshop followed up on the EU Conference on Endocrine Disruptors organised by the European Commission, Directorate-General for Environment in close collaboration with the JRC-IHCP in June 2012.

This workshop sought to lay out the evidence for low dose effects and non-monotonic dose responses in relation to endocrine active chemicals, with the goal of establishing whether the current observations are sufficient to re-examine the ways in which chemicals are tested for endocrine disrupting properties and how risk to human health may be managed.

Most of the participants were in agreement that non-monotonic dose responses do occur and may be expected at some dose ranges for some substances, but the extent to which they might occur at so-called 'low doses' was considered to be a separate issue. There was a suggestion that a definition of 'low dose' would be helpful, as it is currently used with different meanings in different contexts, and that there was also a need to carry out a practical assessment of the type of effects that may be considered adverse, in the context of endocrine disruption. How current test guidelines might be augmented and evaluated with respect to detecting low-dose effects or non-monotonic dose responses was also discussed.

Some recommendations included developing guidance on minimum information requirements for publishing studies investigating endocrine disrupting activity, finding mechanisms for sharing raw data from experimental studies, and creating a knowledge base for compiling findings of non-monotonic dose response relationships.

A report on the workshop is under preparation.

A.6 World Health Organization and United Nations Environment Programme

On 19 February 2013, the United Nations Environment Programme (WHO/UNEP) and World Health Organization (WHO) published a review of the 'State of the Science of Endocrine Disrupting Chemicals – 2012' (WHO/UNEP, 2013). This report updates the IPCS '*Global Assessment of the State-of-the-Science of Endocrine Disruptors'* (WHO/IPCS, 2002) published just over 10 years ago. The new review describes the current global status of scientific knowledge on exposure to, and effects of EDs, and identifies key concerns.

Three strands of evidence fuel concerns over EDs:

- The high incidence and the increasing trends of many endocrine-related disorders in humans, e.g. large proportions (up to 40%) of young men in some countries have low semen quality, which reduces their ability to father children.
- Observations of endocrine-related effects in wildlife populations that have been affected by endocrine disruption, with negative impacts on growth and reproduction.
- The identification of chemicals with endocrine disrupting properties linked to disease outcomes in laboratory studies.

Close to 800 chemicals are known or suspected to be capable of interfering with hormone receptors (some are known to interact with multiple hormone receptors simultaneously), hormone synthesis or hormone conversion. Numerous laboratory studies support the idea that chemical exposures contribute to endocrine disorders in humans and wildlife. The authors consider that EDs may produce non-linear dose response curves both *in vitro* and *in vivo*, by a variety of mechanisms.

The most sensitive window of exposure to EDs is during critical periods of humans and wildlife development, such as during early development. Developmental effects will occur at lower doses than are required for effects in adults, these effects will often be irreversible and may not become evident until later in life. Hence testing for endocrine disruption should encompass the developmental period and include life-long follow-up to assess latent effects.

Human and wildlife populations all over the world are exposed to multiple EDs at the same time. However, only a small fraction of these chemicals have been investigated in tests capable of identifying overt endocrine effects in intact organisms. Significant knowledge gaps exist as to associations between exposures to EDs and other endocrine diseases. This lack of data introduces significant uncertainties about the true extent of risks from chemicals that potentially could disrupt the endocrine system. Disease risk due to EDs may therefore be significantly underestimated.

The authors of the report underline the need for better information on how and when EDs act to reduce exposures during development and prevent diseases from occurring. Endocrine disruption represents a special form of toxicity that should be taken into account when designing studies and interpreting results:

- The effects of the mixtures of chemicals to which humans and wildlife are exposed should be better understood
- The characteristics of the endocrine system that is being disrupted should be taken into account. Endocrine disruption is no longer limited to oestrogenic, androgenic and thyroid pathways and a better understanding of the endocrine systems is therefore needed.
- Testing protocols should cover aspects such as sensitive windows of exposure across the lifespan, low dose effects and NMDRCs.

Internationally agreed and validated test methods for the identification of EDs capture only a limited range of the known spectrum of endocrine disrupting effects. This increases the likelihood that harmful effects in humans and wildlife are being overlooked. The authors recommend the development of WoE approaches to allow for effective consideration of information from all levels (from *in vitro* mechanistic data to human epidemiological data). The need for a transparent methodology for evaluating the strength of evidence of associations between exposure to chemicals and adverse health outcomes is also stressed.

The report underlines the need to reduce exposures to EDs. Government actions to reduce exposures, while limited, have proven to be effective in specific cases (e.g. bans and restrictions on lead, chlorpyrifos, tributyltin, PCBs and some other POPs). This has contributed to decreases in the frequency of disorders in humans and wildlife.

Despite substantial advances in the understanding of EDs in the last 10 years, uncertainties and knowledge gaps still exist that are too important to ignore. The need for an integrated, coordinated international effort to define the role of EDs in current declines in human and wildlife health and in wildlife populations is repeated by the authors.

A.7 EU Member States

In May 2011, the Environmental Protection Agency of the Danish Ministry of the Environment issued a position paper regarding the establishment of criteria for EDs and options for regulation[23]. The report provides a proposal for scientific criteria for the identification of substances with endocrine disrupting properties for humans and the environment. A number of issues relevant for the development of criteria for EDs are discussed and include definition of ED, specificity and level of evidence. The criteria include 3 groups, i.e. ED (group 1), suspected ED (group 2a) and indicated ED (group 2b). The evidence relevant for the 3 groups is discussed and described based on the OECD test methods including the OECD CF. Moreover, some theoretical examples illustrate the use of the criteria. Furthermore, the regulatory use of these criteria in relation to REACH article 57(f) and the new Plant Protection Products (PPP) Regulation is considered. It is proposed that EDs in group 1 should be identified as SVHC in REACH article 57/f) and as ED substances under PPP. For suspected and

[23] Available at: http://www.mst.dk/English/Chemicals/endocrine_disruptors/danish_proposal_for_criteria_for_endocrine_disruptors_submitted_to_the_EU/

indicated EDs (group 2a and 2b), further data may be necessary to evaluate whether the substances is an ED (group 1).

In May 2011, the UK Health and Safety's Chemicals Regulation Directorate and the German Institute for Risk Assessment issued a position paper on the regulatory definition of an ED in relation to potential threat to human health[24]. The document proposes that an exogenous substance or mixture is regarded as an ED of very high regulatory concern if it alters function(s) of the endocrine system and consequently causes adverse effects in an intact organism, or its progeny, or (sub)populations. In addition, i) the adverse effects should have been seen in one or more toxicity studies of acceptable quality, in which the substance was administered by a route relevant for human exposure, ii) a plausible mode-of-action/mechanistic link between the toxic effects of concern and endocrine disruption, iii) the effects seen in experimental animals should be of potential relevance to human health and iv) serious adverse effect(s) in animal studies related to endocrine disruption should have been produced at a dose at or below the relevant guidance value for the application of Category 1 'Specific Target Organ Toxicity-Repeated Exposure, STOT-RE' classification & labelling.

In March 2012, the French Agency for Food, Environmental and Occupational Health & Safety (ANSES) published an opinion regarding a request from the French Directorate General for Food, Directorate General for Health, Directorate General for Risk Prevention, Directorate General for Competition, Consumer Affairs and Fraud Control, and the Directorate General for Labour for scientific and technical support for the revising of the European strategy on endocrine disruptors[25]. The document proposes to use the WHO/IPCS (2002) definitions for endocrine/potential EDs. The ANSES's proposal for the scientific criteria for identifying an ED that are applicable to the REACH, Biocides and PPP Regulations is based on that of the Danish Authorities which separates the endocrine disrupting substances into two categories, confirmed and potential, with this second category being further divided into two sub-categories. ANSES proposes adding to the Danish position a regulatory criterion from the joint proposal by the United Kingdom and German authorities: a limit applicable only to plant protection substances. Substances with harmful endocrine disrupting effects, observed in mammals at a dose lower than 10 mg/kg bw/day, should be placed in Category 1 and thus cannot be approved under the Regulation. This value, established to facilitate decision-making, is based on the Regulation (EC) No 1272/2008 on classification, labelling and packaging of substances and mixtures (STOT-RE 1 effect).

In May 2012, the Danish Centre on Endocrine Disrupters evaluated 22 substances from the 'Substitute It Now' (SIN) List 2.0 identified by the Non-Governmental Organisation ChemSec as substances of very high concern[26]. The substances were categorised on the basis of the Danish proposal for criteria for EDs (see above), and according to the criteria proposed in the above-mentioned joint British-German position paper that proposes the use of a potency cut-off criteria. The evaluation using the Danish criteria lead in most cases to the same conclusion as the SIN List evaluation, with 15 of the 22 substances categorised as EDs (category 1). Six of the 7 remaining compounds were categorised as suspected EDs (category 2A). The use of the potency cut-of criteria leads to 4 of the 15 ED substances (category 1) being considered as substances of very high concern.

In February 2013, the Swedish Chemicals Agency published a position paper on the possibility to determine threshold levels for EDs; the paper is based on a review of 15 publications on endocrine disruption. The authors concluded that an ED is identified if a plausible link between the endocrine mode of action and the adverse effect can be demonstrated, and that the current available standardised test methods are limited to EATS hormonal systems.

[24] Available at: http://www.bfr.bund.de/cm/343/regulatory_definition_of_an_endocrine_disrupter_in_relation_to_potential_threat_to_human_health.pdf
[25] Available at: http://www.anses.fr/Documents/DPR2012sa0033EN.pdf
[26] Available at: http://www.mst.dk/NR/rdonlyres/CDA4EB4F-1554-4754-A0F9-77D73BCA0228/0/SINreportandAnnex.pdf

The authors of the paper considered more specifically whether the scope of Article 60(3) of the REACH Regulation (Regulation (EC) No 1907/2006) should be extended to include EDs, and whether EDs should be considered like CMR substances for which it is not possible to determine a threshold. The performed literature review provided arguments both for and against assuming a threshold for EDs. The authors conclude that the decision to accept or not a non-threshold model has to be based on the endpoint under consideration and what is known about its mode or mechanism of action. "*Hence, the assumption of no threshold may be as valid, or questionable, for EDs as for genotoxic carcinogens*". The authors consider that it will not be possible to arrive at a robust, reliable and sufficiently protective threshold values for EDs in the near future, and therefore recommend that EDs are covered by Article 60(3) of the REACH Regulation.

A.8 Stakeholders' initiatives

PAN Europe

In its position paper on criteria for endocrine disrupting pesticides, published in May 2011 (PAN Europe, 2011), PAN Europe proposes that a substance should be considered as having endocrine disrupting properties when effects on the endocrine system are observed, including effects secondary to other toxic effects. A known mechanism of action is not necessary. In order to identify substances having endocrine disrupting properties, it is necessary to study all hormonal systems, perform low-dose testing, consider the notion of exposure window and therefore administer the substance to animals during their development. The notion of threshold should not be used for endocrine disrupting properties because of available examples of non-linear dose responses.

Possible definitions EDs are not discussed, as there is only a need to identify criteria for endocrine disrupting properties. PAN Europe proposes that the approach should be based on the hazard and not on risk assessment.

An in-depth review of the scientific literature should be undertaken for a hazard assessment of the studied substance, with preference to be given to data from independent organisations. Concerning testing of substances, PAN Europe recommends a modern study protocol to be developed by independent scientists working on endocrine disruption.

Regarding the interpretation of study results, the effects observed in animals should by default be considered relevant for humans. If there is doubt about the adverse effects of chemicals with endocrine disrupting properties, the precautionary principle must be used and the chemical withdrawn from the market until further studies are evaluated.

CHEM Trust

In September 2011, CHEM Trust provided a "Contribution to the Ongoing Debate on Criteria for endocrine disrupting compounds" (CHEM Trust, 2011a); the paper was developed with input from WWF European Policy Office. In April 2012, CHEM Trust and HEAL discussed in a joint briefing "Challenges and solutions in the regulation of chemicals with endocrine disrupting properties (CHEM Trust, 2011b). The following principles were recommended to play a role when developing an assessment and identification scheme for EDs:

Concerning the definition of an ED, CHEM Trust notes that the WHO/IPCS (WHO/IPCS, 2002) definition provides a useful scientific working definition but requires too high a bar of proof that a substance, by disruption of the endocrine system, 'consequently causes adverse health effects'. CHEM Trust underlines the difficulty to establish the mechanism of action and therefore favours the definition put forward by experts at the Danish Centre on Endocrine Disrupters which still requires adverse effects but removes the word 'consequently'.

Based on the above-suggested definition, criteria for the identification of an ED should address whether the chemical (i) has ED properties (exact knowledge about the mechanism of action of how the endocrine disrupting property is exerted should not be a prerequisite) and (ii) has the ability to cause adverse effects (in laboratory test systems). Regarding the adversity criteria, CHEM Trust underlines that regulations refer to 'probable serious effects' and 'may cause adverse effects'; therefore adverse effect should be reasonably predicted. A potency threshold should not be included in the criteria to identify chemicals with ED properties

CHEM Trust underlines that testing requirements for chemicals should be improved, so that these are better orientated to identifying EDs. Given the limitations of current OECD test methods, non-OECD test methods should be given due weight in hazard assessment, and *in vitro* information should be used as supporting evidence to avoid unnecessary testing.

CHEM Trust acknowledges other ED properties, such as potential non-linear dose response curves and low dose effects, including possible lack of thresholds.

With regard to risk management considerations, CHEM Trust recommends that a chemical with endocrine disrupting properties should trigger regulation even if these are not the 'lead' effect. Potency thresholds should not be used to exclude certain chemicals with endocrine disrupting properties from stricter regulation, as it would unfortunately result in the legislation not achieving its goal of protecting health.

ECETOC

The ECETOC position on criteria for evaluating and identifying EDs starts with the Weybridge definition of an ED:

> *"An ED is an exogenous substance that causes adverse health effects in an intact organism, or its progeny, secondary to changes in endocrine function"*. (EC, 1997)

However, as noted by Bars et al. (2011), definitions from WHO/IPCS (2002), EC (1999), and Japanese Ministry of Environment (2005) are equally relevant, in sharing a common element, that of adversity of the effect. The ECETOC approach for identification of EDs thereby holds as central the need to demonstrate both an endocrine-active mechanism of action of a chemical AND a significant effect on an apical endpoint relevant to the protection goals (ECETOC, 2009). In the case of ecotoxicity testing therefore, apical effects would need to be considered as population relevant e.g. affecting survival or reproduction. ECETOC recommends using a WoE assessment technique to evaluate the range of (eco)toxicity data, including *in vitro* and *in vivo* targeted studies, supporting studies and multi-endpoint apical studies available.

A testing strategy that combines whole organism regulatory tests, with information on mode of action is proposed. Tests are categorised as being either:

- *in vitro* screens targeted to a particular mode of action:
    - ER/AR binding assays, ER Stably Transfected Transcriptional Activation (STTA) assay, aromatase activity assay and H295R screen for steroidogenic activity
- mechanistically informative *in vivo* studies targeted to one or more modes of ED activity
    - mammal
        - Uterotrophic (OECD TG 440) and Hershberger (OECD TG 441) assays
    - Ecotox
        - fish 21d screening assay (OECD TG 230)
        - fish short-term reproduction assay (OECD TG 229)
        - Amphibian metamorphosis assay (OECD TG 231)
- definitive (apical) and supporting *in vivo* studies that incorporate multiple endpoints that can be indicative of adverse effects and can be used in wider risk assessment

- mammal:
  - Pubertal male/female rat assays, US EPA, (U.S. EPA, 2011d, 2011e)
  - enhanced 28d repeat dose oral toxicity study (OECD TG 407)
  - sub chronic studies; OECD TG 408; TG 409, (OECD, 1998)
  - pre-natal development toxicity study (OECD TG 414), chronic/carcinogenicity studies (OECD TGs 451, 452, 453)
  - single generation rodent study (OECD TG 415 and extended version)
  - rodent 2-generation study (OECD TG 416)
- Ecotox:
  - Fish sexual development test (OECD TG 234)
  - modified fish full cycle test; e.g. modification of US EPA OPPTS 850.1500, (U.S. EPA, 2004)
  - Avian and amphibian partial or full life-cycle tests still in development

A decision matrix for evaluating outcomes from tests in these categories has been developed (ECETOC, 2009). This combines test outcomes from apical definitive/supporting tests (indicative of an adverse effect) and mechanistically informative *in vivo* or *in vitro* assays. Only when testing indicates both 'adverse effects giving concern for endocrine toxicity' from apical/supporting studies and evidence of 'endocrine activity giving concern for endocrine toxicity' are the data considered to provide 'Sufficient evidence of ED as per the Weybridge definition'. In the absence of convincing evidence for the endocrine mode of action, the default is to assume human relevance. A series of supplemental decision trees are provided for evaluating such data relevant to mammal studies for human health, and for wildlife species (fish and amphibians or birds and mammals). In cases where there is deemed to be sufficient evidence of ED as per Weybridge there is further consideration of the toxicity profile of the chemical to consider:

- Are the effects specific? (i.e. can the effects be attributed to indirect effects arising from systemic toxicity). The concept of lead toxicity vs. specificity of endocrine-mediated effects should be taken into account. The acceptable degree of separation between the lead effect and the endocrine-mediated effect should be assessed on a case-by-case basis; a factor of 10 was suggested; i.e. if the degree of separation is >10X the substance should not be considered as an ED of concern. If the degree of separation is <10X, then the substance should be considered as an ED of concern. For ecotoxicological assessment, a greater degree of separation between the lowest lead effect and the endocrine-mediated effect would be required for the aquatic environment than the terrestrial environment due to the higher diversity of species
- Is the ED mechanism of action relevant to human health/environmental species? (unless exposure is negligible, reflecting the caveat in section 3.8.2 of Regulation (EC) No 1107/2009)
- What is the potency of ED? The concept of potency could serve to discriminate the substances of higher concern from those of lesser concern, although it is seen as a poor substitute for risk assessment. Substances with ED properties which are not caught by the potency assessment and possible cut-off criterion are still considered as EASs and should be subject to a standard risk assessment. ECETOC proposes the following criteria when considering ED relative to human health concerns: 1) dose/concentration 2) exposure duration (acknowledging that weak EDs may only manifest detectable effects after longer exposure durations 3) type and severity of endocrine effects 4) number of species affected (acknowledging potential for read-across between species/ecotox-human health models). For evaluating the 'potency' of EDs with regard to environmental effects, ECETOC proposes to consider the same criteria as for human health concerns with the addition of 'specificity of endocrine effects in relation to other taxonomic groups (e.g. comparing potency of ED activity in a fish with other non-ED effects in algae).

B. OVERVIEW OF PROVISIONS FOR ENDOCRINE DISRUPTORS IN EUROPEAN LEGISLATION

Regulation (EC) No 1907/2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH)

Substances having endocrine disrupting properties may be identified as Substances of Very High Concern (SVHC) and may subsequently be made subject to the authorisation procedure under the European Union REACH Regulation (EC) No 1907/2006.

Substances may be identified as SVHC under Article 57 of the REACH Regulation if they have certain properties, i.e. substances meeting the criteria for classification as Carcinogenic, Mutagenic or Toxic for Reproduction (CMR), Category 1A or 1B in accordance with the CLP Regulation[27]; PBT substances, vPvB substances as defined by Annex XIII of the REACH Regulation; and substances for which there is scientific evidence of probable serious effects to human health or the environment which give rise to an equivalent level of concern to the aforementioned substances. These latter Art. 57(f) substances may include EDs and shall be identified on a case-by-case basis. Authorisation requirements apply to SVHC that are included in Annex XIV of REACH.

If the substance is initially identified as an SVHC and subsequently subjected to authorisation, then any manufacturer, importer or downstream user must not, after a certain date, place that substance on the market for a use, or use it itself without prior authorisation, unless the use has been exempted. There is no tonnage limit for the authorisation requirement. An authorisation may be granted if it has been shown that the risks arising from the use of the substance are adequately controlled or there are no suitable alternative substances or technologies and the socio-economic benefits outweigh the risk(s) identified.

The guidance document that was written for the competent authorities for the preparation of a dossier on the identification of SVHC will be revised in line with the outcome of the European Commission's upcoming proposal for criteria for the identification of substances with endocrine disrupting properties (the proposal is planned for end of 2013).

Regulation (EC) No 1107/2009 concerning the placing of plant protection products on the market

Substances regarded as having endocrine disrupting properties that may be harmful to humans or non-target organisms, unless the exposure is negligible under the conditions of use, cannot be authorised pursuant to Regulation (EC) No 1107/2009. Article 4(7) indicates the conditions and situations that might justify an exemption and also the measures to mitigate the risks, while also informing the European Commission. Moreover, by no later than 14 December 2013, the Commission is required to present proposed measures concerning specific scientific criteria for the determination of endocrine disrupting properties (Annex II, Point 3.6.5 of Regulation (EC) No 1107/2009).

However, pending the adoption of these criteria, substances that are or have to be classified, pursuant to the provisions of Regulation (EC) No 1272/2008 (CLP Reg.), as Carcinogenic Category 2 or Toxic for Reproduction Category 2, shall be considered as having endocrine disrupting properties.

In addition, substances such as those that are or have to be classified, pursuant to the provisions of the CLP Regulation, as Toxic for Reproduction Category 2 and which have toxic effects on the endocrine organs may be considered as having such endocrine disrupting properties.

Regulation (EU) No 528/2012 concerning the making available on the market and use of biocidal products

Biocidal substances are not approved if they have endocrine disrupting properties. They are identified using the criteria described in Article 57 (f) of the REACH Regulation. This non-approval does not

[27] Regulation (EC) No 1272/2008 on classification, labelling and packaging of substances and mixtures

apply if the risk to humans and the environment is negligible, if the substance is essential to combat a serious health risk, or if such non-approval would result in disproportionate negative impacts on society relative to the risks to humans and the environment.

Scientific criteria for the determination of endocrine disrupting properties must be adopted no later than 13 December 2013. In the meantime, endocrine disrupting substances are considered to be those substances which, under the provisions of Regulation (EC) No 1272/2008, are - or should be - classified as:

- carcinogenic category 2 and toxic for reproduction category 2;
- toxic for reproduction category 2 and which have toxic effects on the endocrine organs;
- or substances that have been identified as having endocrine disrupting properties under Articles 57 (f) and 59 (1) of Regulation (EC) No 1907/2006.

Regulation (EC) No 1223/2009 on cosmetic products

Endocrine disrupting substances are currently not restricted in the scope of Regulation (EC) 1223/2009 on cosmetic products. The Regulation should be reviewed with regard to substances with endocrine-disrupting properties when Community or internationally agreed criteria for identifying substances with endocrine-disrupting properties are available, or at the latest on 11 January 2015.

Directive 2000/60/EC establishing a framework for Community action in the field of water policy

The Water Framework Directive (WFD) sets environmental objectives of good chemical status for surface waters and for the prevention of pollution of groundwater.

At the national level, Member States are required to identify chemical pollutants of relevance for each of the water bodies, to set quality standards for water, to establish emission control measures and to achieve these standards by 2015. An indicative list of the main pollutants is included in Annex VIII of the Directive; a specific category includes those 'substances and preparations, or the breakdown products of such, which have proved to possess carcinogenic or mutagenic properties or properties which may affect steroidgenic, thyroid, reproduction or other endocrine-related functions in or via the aquatic environment' (Annex VIII – Group 4).

At the Community level, the WFD sets out a strategy against pollution of surface waters by chemical pollutants (Article 16). This strategy includes the identification of substances of particular concern at Community level. A first list of 33 substances was adopted in 2001[28]; of these 33 substances, 21 are candidate endocrine disrupting substances for which evidence of endocrine disruption or potential endocrine disruption was found in the BKH Consulting Engineers (Delft, The Netherlands) reports (2000-2003). This first list was replaced by Annex II of the Directive on Environmental Quality Standards (Directive 2008/105/EC) (EQSD)[29], also known as the Priority Substances Directive. As required by the WFD and EQSD, the Commission subsequently reviewed the list and in 2012 it put forward a proposal for a Directive amending the WFD and the EQSD as regards priority substances. This proposal[30] includes a revised (second) list of priority substances, and provisions to improve the functioning of the legislation. Although no direct reference is made to ED properties of these substances, endocrine disruption could become an important criterion for sorting substances or groups of substances into this group.

[28] Decision No 2455/2001/EC of the European Parliament and of the Council of 20 November 2001 establishing the list of priority substances in the field of water policy and amending Directive 2000/60/EC

[29] See http://ec.europa.eu/environment/water/water-dangersub/pri_substances.htm#dir_prior

[30] See COM(2011)876 available at http://ec.europa.eu/environment/water/water-dangersub/lib_pri_substances.htm#prop_2011_docs

C. OVERVIEW OF OECD GUIDANCE, TEST GUIDELINES AND ONGOING ACTIVITIES

The OECD work on EASs is carried out by its working groups of the Environment Health and Safety Testing Assessment programme, as a special activity supervised by the EDTA6, and reporting to the WNT. Three further expert groups manage the work: the Validation Management Groups on mammals (VMG – Mamm), on ecotoxicology (VMG-eco) and on non-animal (VMG-NA) testing - 3Rs. The Endocrine Disrupters Testing and Assessment Task Force (EDTA-TF) was created in 2002, and adopted a Conceptual Framework (CF) for Testing and Assessment of Endocrine Disrupting Chemicals (i.e. a toolbox with screening and test methods for oestrogen-, androgen- and thyroid-mediated modes of action). An overview with 'Information on OECD Work related to Endocrine disruptors' is given in OECD (2012f).

C.1 The OECD Guidance Document of August 2012

The OECD Guidance Document of August 2012 is a non-binding guidance on standardised test guidelines for evaluating chemicals for endocrine disruption (OECD, 2012a) incorporating the revised CF for testing and assessment of endocrine disruptors, as the main output of the EDTA6. This revised CF lays out possible screens, tests and data sources for EATS EDs, organised into 5 Levels which range from Level 1 (desirable and non test information), through Level 2 (*in vitro* assays providing data about selected endocrine mechanism(s)/pathway(s)), Level 3 (*in vivo* assays providing data about selected endocrine mechanism(s)/pathway(s)), Level 4 (*in vivo* assays providing data on adverse effects on endocrine relevant endpoints), to Level 5 (*in vivo* assays providing more comprehensive data on adverse effects on endocrine relevant endpoints over extensive parts of the lifecycle of the organisms). Levels 3-5 present mammalian and non-mammalian assays in two separate columns, and the assays comprise both existing OECD test guidelines (TGs) and TGs in development and validation at OECD. See extract from the guidance here below (i.e. from OECD GD 150, OECD, 2012a, pp. 385-387).

Testing strategies

The CF attempts to address ED testing by working as a tool kit, rather than a sequential testing programme, and by defining whether information that can be generated in a particular method is mechanistically informative, apical (and therefore of relevance to adversity), or both. The CF is therefore not intended as a testing strategy and does not include evaluation of exposure.

On the other hand, the OECD held an international expert workshop in 2010 which reviewed available and forthcoming OECD test guidelines for measuring toxicity to fish, and *inter alia* developed an integrated testing strategy that included the new fish tests sensitive to endocrine disrupting substances and at the same time minimised the use of fish in chemical testing programmes. After revision and endorsement by OECD member states, this strategy has now been published (OECD, 2012g).

A summary of the proposed generic testing strategy is shown in Fig. 1. When using the flow chart, it is essential to read it in conjunction with the accompanying text in (OECD, 2012g), but the main points regarding EDs will be briefly described below. The underlying philosophy is to use a WoE approach when deciding how to proceed at certain points. The strategy starts with an intensive data-gathering and problem formulation step, including the evaluation of a range of physico-chemical and fate information, toxicity predictions from (Q)SARs, read-across data from other chemicals and organisms

(including mammals), and consideration of any existing *in vitro* or *in vivo* toxicity data (including any relevant information on potential endocrine disruption).

The user is required to consider whether any suspicions regarding endocrine disruption have been raised following the initial data gathering, read-across and *in vitro* screening. If no suspicions exist, one proceeds down the standard route of considering the need for a fish early life stage (ELS) test; OECD TG 210 (OECD, 1992) or similar. The alternative is for the user to consider various forms of EDs screening or testing, depending on the strength of existing data. Weak suspicions trigger *in vivo* fish screening using OECD TG 229 or 230, while moderate suspicions trigger the fish sexual development test (FSDT – OECD TG 234) or a fish partial life cycle reproduction test. Strong suspicions of endocrine activity will generally result in a requirement for fish full life cycle (FFLC) testing or multi-generation (MG) testing. Any alerts about endocrine disruptive activity resulting from the *in vivo* screens, FSDT or partial life cycle reproduction tests will in turn trigger FFLC or MG testing. On the other hand, if the EDS-sensitive screens do not raise concerns, the user is then required to consider ELS testing, which may in turn lead on to FFLC testing in some circumstances.

Interpretation of results

In the OECD guidance document, scenarios are given that represent all the possibilities of positive or negative results in combination with the presence of existing data and how to interpret such scenarios. For each assay under consideration, the GD considered 16 hypothetical scenarios, ranging from very data-rich situations in which positive, negative or equivocal data are available from both existing *in vitro* and *in vivo* assays, to very data-poor examples where little or no existing data are available. Half the scenarios cover situations in which the assay in question has provided a positive result (i.e. it has indicated that the substance has produced a significant response), and half cover negative assay responses. For those assays which provide both mechanistic and apical information (e.g. vitellogenin and fecundity data in TG 229), the scenarios are elaborated even further to distinguish between assays providing different combinations of these data categories. In all cases, the interpretational advice is based on a WoE assessment in the light of the existing data posited in the 16 scenarios.

Example of guidance on interpretation of assay results for birds:

Together with the results of existing *in vitro* and *in vivo* information (see information in section 4.2.3) and the results of the OECD TG 206 and the ATGT (see section 4.2.4) the potential hazards of endocrine disrupting substances to birds is summarised in the below Table 4 (extracted from OECD GD 150, OECD, 2012a).

The GD also provides advice on a possible next testing step which could be taken if the user considers it necessary to lessen uncertainty about the test substance. It should, however, be noted that the GD was not intended to provide a comprehensive hazard testing scheme.

The OECD is committed to update this guidance with further evaluation of screens and tests with sensitivity for EDSs, and the 3 case studies (OECD, 2012c) did not uncover any significant problems, although further studies with negative substances would be desirable.

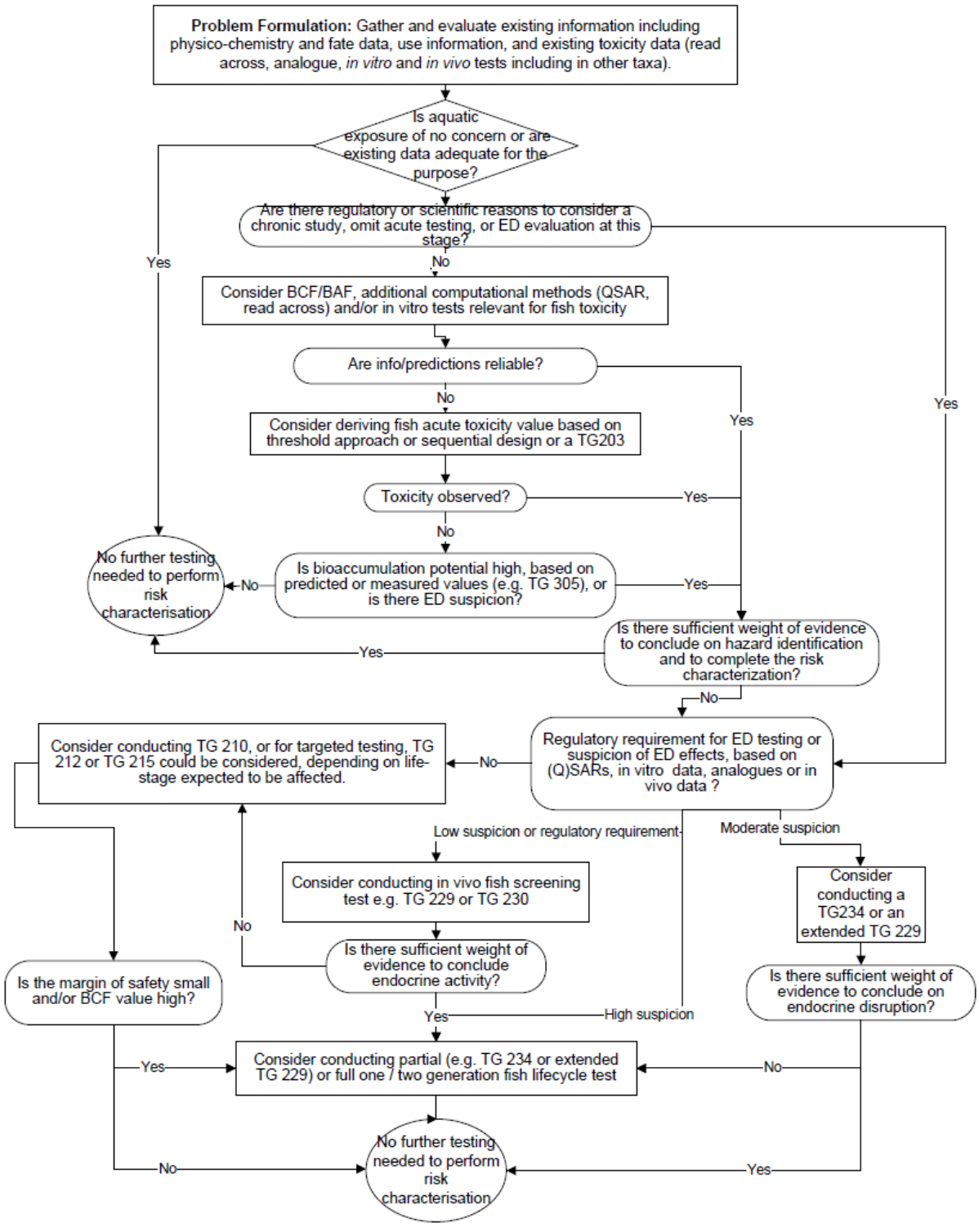

Figure 1: Generic fish testing strategy. Reproduced with permission from OECD (2012g) Fish Toxicity Testing Framework. Series on Testing and Assessment no. 171, ENV/JM/MONO(2012)16 http://search.oecd.org/officialdocuments/displaydocumentpdf/?cote=ENV/JM/MONO%282012%2916&doclanguage=en.

Table 4: Possible conclusion from results of the TG 206 or ATGT, *in vitro* and *in vivo* assays. Presence of positive (+), negative (-) or equivocal/absent (Eq/0) existing results.

Scenario	*Results*				**Tentative conclusions**	
	***in vitro* mechanistic data**	***in vivo* effects of concern**	**TG 206**	**ATGT**	**Based on TG 206**	**Based on ATGT**
A	+	+	+	+	Probably an ED	Strong evidence for adverse effects
B	+	-	+	+	Probably an ED	Strong evidence for adverse effects
C	+	Eq/0	+	+	Probably an ED	Strong evidence for adverse effects
D	-	+	+	+	May be an ED	Strong to medium strong evidence
E	-	-	+	+	Unlikely an ED	Strong to medium strong evidence
F	-	Eq/0	+	+	Unlikely an ED	Strong to medium strong evidence
G	Eq/0	+	+	+	May be an ED	Strong to medium strong evidence
H	Eq/0	-	+	+	May be an ED	Strong to medium strong evidence
I	Eq/0	Eq/0	+	+	May be an ED	Strong to medium strong evidence
J	+	+	-	-	Probably not an ED in birds	Probably not an ED in birds
K	+	-	-	-	Probably not an ED in birds	Probably not an ED in birds
L	+	Eq/0	-	-	Probably not an ED in birds	Probably not an ED in birds
M	-	+	-	-	Probably not an ED in birds	Probably not an ED in birds
N	-	-	-	-	Probably not an ED in birds	Probably not an ED in birds
O	-	Eq/0	-	-	Probably not an ED in birds	Probably not an ED in birds
P	Eq/0	+	-	-	Probably not an ED in birds	Probably not an ED in birds
Q	Eq/0	-	-	-	Probably not an ED in birds	Probably not an ED in birds
R	Eq/0	Eq/0	-	-	Probably not an ED in birds	Probably not an ED in birds

Extract of the CF list

Conceptual Framework for Testing and Assessment of Endocrine Disrupters, reproduced with permission from the OECD (2012a) Guidance document on standardised test guidelines for evaluating chemicals for endocrine disruption. Series on Testing and Assessment no. 150, ENV/JM/MONO(2012)22 http://search.oecd.org/officialdocuments/displaydocumentpdf/?cote=env/jm/mono%282012%2922&doclanguage=en (OECD, 2012a), pp. 385-387

The Conceptual Framework lists the OECD TGs and standardized test methods available, under development or proposed that can be used to evaluate chemicals for endocrine disruption. The Conceptual Framework is intended to provide a guide to the tests available which can provide information for endocrine disruptors assessment but is not intended to be a testing strategy. Furthermore, this Conceptual Framework does not include evaluation of exposure, however this should be included when deciding whether further testing is needed. Further information regarding the use and interpretation of these tests is available in GD 150 *(i.e.* this GD).

Mammalian and non mammalian Toxicology

Level 1: Existing data and non-test information

- Physical & chemical properties, e.g., MW reactivity, volatility, biodegradability
- All available (eco)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) (Mammalian and non mammalian methods)

- Estrogen or androgen receptor binding affinity
- Estrogen receptor transactivation (OECD TG 455)[31]
- Androgen or thyroid transactivation (If/when TGs are available)
- Steroidogenesis *in vitro* (OECD TG 456)
- MCF-7 cell proliferation assays (ER ant/agonist)
- Other assays as appropriate

[31] OECD Test Guideline 457: 'BG1Luc Estrogen Receptor Transactivation Test Method for Identifying Estrogen Receptor Agonists and Antagonists'. New test guideline added in November 2012.

Level 3: *In vivo* assays providing data about selected endocrine mechanism(s) / pathway(s)[1]

Mammalian Toxicology for Level 3

- Uterotrophic assay (OECD TG 440)
- Hershberger assay (OECD TG 441)

Non-Mammalian Toxicology for Level 3

- Xenopus embryo thyroid signalling assay (When/if TG is available)
- Amphibian metamorphosis assay (OECD TG 231)
- Fish reproductive screening assay (OECD TG 229)
- Fish screening assay (OECD TG 230)
- Androgenized female stickleback screen (GD 140)

Level 4: *In vivo* assays providing data on adverse effects on endocrine relevant endpoints [2]

Mammalian Toxicology for Level 4

- Repeated dose 28-day study (OECD TG 407)
- Repeated dose 90-day study (OECD TG 408)
- 1-generation reproduction toxicity study (OECD TG 415)
- Male pubertal assay (see GD 150 Chapter C4.3)[3]
- Female pubertal assay (see GD 150 Chapter C4.4)[3]
- Intact adult male endocrine screening assay (see GD 150 Chapter Annex 2.5)
- Prenatal developmental toxicity study (OECD TG 414)
- Chronic toxicity and carcinogenicity studies (OECD TG 451-3)
- Reproductive screening test (OECD TG 421 if enhanced)
- Combined 28-day/reproductive screening assay (OECD TG 422 if enhanced)
- Developmental neurotoxicity (OECD TG 426)

Non-Mammalian Toxicology for Level 4

- Fish sexual development test (Draft OECD TG 234)
- Fish reproduction Partial Lifecycle Test (when/If TG is Available)
- Larval amphibian growth & development assay (when TG is available)
- Avian reproduction assay (OECD TG 206)
- Mollusc partial lifecycle assays (when TG is available)[4]
- Chironomid toxicity test (TG 218-219)[4]
- Daphnia reproduction test (with male induction) (OECD TG 211)[4]
- Earthworm reproduction test (OECD TG 222, 2004)[4]
- Enchytraeid reproduction test (OECD TG 220, 2004)[4]
- Sediment water lumbriculus toxicity test using spiked sediment (OECD TG 225, 2007)[4]
- Predatory mite reproduction test in soil (OECD TG 226, 2008)[4]
- Collembolan reproduction test in soil (TG OECD 232, 2009) [4]

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 [2]

Mammalian Toxicology for Level 5

- Extended one-generation reproductive toxicity study (OECD TG 443)[5]
- 2-Generation reproduction toxicity study (OECD TG 416 most recent update)

Non-Mammalian Toxicology for Level 5

- Fish lifecycle toxicity test (FLCTT) (when TG is available)
- Medaka multigeneration test (MMGT) (when TG is available)
- Avian 2 generation reproductive toxicity assay (when TG is available)
- Mysid lifecycle toxicity test (when TG is available)[4]
- Copepod reproduction and development test (when TG is available)[4]
- Sediment water chironomid life cycle toxicity test (OECD TG 233)[4]
- Mollusc full lifecycle assays (when TG is available)[4]
- Daphnia multigeneration assay (if TG is available) [4]

[1] Some assays may also provide some evidence of adverse effects.

[2] Effects can be sensitive to more than one mechanism and may be due to non-ED mechanisms.

[3] Depending on the guideline/protocol used, the fact that a substance may interact with a hormone system in these assays does not necessarily mean that when the substance is used it will cause adverse effects in humans or ecological systems.

[4] At present, the available invertebrate assays solely involve apical endpoints which are able to respond to some endocrine disruptors and some non-EDs. Those in Level 4 are partial lifecycle tests, while those in Level 5 are full- or multiple lifecycle tests.

[5] The new EOGRT study (OECD TG 443) is preferable for detecting endocrine disruption because it provides an evaluation of a number of endocrine endpoints in the juvenile and adult F1, which are not included in the 2-generation study (OECD TG 416) adopted in 2001

Notes to the OECD Revised Conceptual Framework

Note 1: Entering at all levels and exiting at all levels is possible and depends upon the nature of existing information and needs for testing and assessment.

Note 2: The assessment of each chemical should be made on a case-by-case basis, taking into account all available information.

Note 3: The framework should not be considered as all inclusive at the present time. At levels 2, 3, 4 and 5 it includes assays that are either available or for which validation is under way. With respect to the latter, these are provisionally included.

C.2 Ongoing activities at OECD

Extensive international OECD collaborative efforts have been conducted previously, to improve the endocrine relevance of tests contained within the OECD CF, at review, guidance and validation levels. Thus for example, in addition to the development of *in vitro* tests (oestrogen, androgen receptor and steroidogenesis) in Level 2 of the CF, and how metabolism could be better incorporated, TGs have also been examined for ways in which endocrine endpoint information could be better addressed and collected. Examples for human health hazard assessment include TG 407, which underwent an additional validation exercise, after some additional endocrine mediated endpoints were added (such as weight of testes, adrenals, prostate, and histopathology of gonads accessory sex organs, uterus, adrenal, thyroid and vagina), and put into context with other toxicological effects. Parameters for which insufficient data were available, or only showed weak evidence of their ability to help in the detection of EDs in the validation exercise were proposed as optional endpoints (e.g. the ovary, uterus and thyroid weights and histopathology of vaginal smears, male mammary glands, and pituitary; also measurement of circulating levels of T3, T4 and TSH). The extended one generation assay reproductive assay (TG 443) was designed on a modular 'trigger' basis specifically to address reproductive/developmental (but also immune and neurotoxicity) endocrine endpoints not covered in the two generation assays. In particular, it addresses reproductive endpoints that require the interaction of males and females, females with conceptus, females with offspring, and the F1 generation until after sexual maturity. Endocrine relevant endpoints include for example the anogenital distance and the presence of nipples/areolae in male pups.

Examples of endocrine relevant non mammalian toxicity TGs and guidance include the amphibian metamorphosis assay (TG 231), the fish screening assay (TG 230), the fish short-term reproduction assay (TG 229) the androgenised female stickleback screen (OECD GD 140), and the fish sexual development test (TG 234).

As the work will continue to deliver more suitable TGs, the following ongoing activities within OECD are relevant for EASs:

VMG-NA has provided TG 456 on steroidogenesis (that is already included in the CF) and is carrying out further related work to support TG development: AOPs, HTS, metabolism, species differences, new endpoints, Ad Hoc QSAR group.

The OECD described in its Work plan of the Test Guideline Programme (OECD, 2012d), the following projects relevant for screening/testing chemicals for endocrine disruption:

Related to health effects and endocrine disruption:

> **Project 4.31: EDTA Activity - New TG: Human Recombinant Estrogen Receptor Alpha Binding Assays (hrERA, 2 protocols).** Lead: United States + European Commission + Germany + Japan. Project status and milestones: Validation plan discussed the meeting of the VMG-NA (2007); Validation (complete) was presented at the meeting of the VMG-NA (2010); Data analysis and compilation completed in 2011; Work ongoing.
>
> **Project 4.33: EDTA Activity - New TG: Stably Transfected Transcriptional Activation (STTA). Assay for the detection of androgenic and anti-androgenic activity of chemicals.** Lead: Japan. Project status and milestones: Draft validation report and draft TG submitted to the Secretariat in 2010; Draft validation report submitted to the VMG-NA in December 2010; Peer review report available in February 2011; Draft peer review report (with draft WNT Statement on the follow-up to the peer review) endorsed/agreed at the 2011 WNT meeting; Discussion of chemicals to be included in an additional validation at the VMG-NA meeting; Work ongoing.
>
> **Project 4.34: EDTA Activity - New TG for a stably Transfected Transactivation (STTA) Assay for the detection of anti-estrogenic activity of chemicals.** Lead: Japan. Project status and milestones: Collection of validation data expected in 2nd quarter 2012; Work ongoing.

Project 4.47: EDTA Activity - TG for an MCF-7 Cell Proliferation Assay for the Detection of Estrogen Receptor Agonist and Antagonist. Lead: United States. Project status and milestones: Information on the added value, scope, and possible place in an endocrine disruptor assessment scheme expected to be provided by the VMG-NA before the first commenting round; Validation study completed in January 2011; Validation report expected end of March 2012; US expected to make a decision on whether a Peer Review will be conducted and a draft test guideline developed.

Project 4.48: EDTA Activity - TG for a Chimpanzee Recombinant Androgen Receptor Binding Assay. Lead: United States. Project status and milestones: Information on the added value, scope, and possible place in an endocrine disruptor assessment scheme expected to be provided by the VMG-NA before the first commenting round; Pre-validation ongoing; Pre-validation report expected when the pre-validation is completed. This project is not expected to be continued after that stage.

****Project 4.55: GD on internal triggers for US/Canada (GD 117) and GD Supporting the TG for an Extended One Generation Reproductive Toxicity Study (draft GD 151).** Lead: Secretariat. Project status and milestones: Preliminary draft GD supporting the Extended One Generation Reproductive Toxicity Study developed in parallel with the draft TG; Draft GD on internal triggers approved at the 2011 WNT meeting; GD published as No 117; Request for WNT comments on the draft GD 151 in September 2011; Discussion of draft GD 151 at an expert meeting on 25-26 January 2012 in Arlington (VA, United States); WNT Comments on a revised draft GD requested in April 2012; GD expected to be approved at the 2013 WNT meeting.

Note: This work on an extended one-generation reproductive toxicity study TG 443 & GD 151 has the objective to evaluate specific life stages not covered by other types of toxicity studies and test for effects that may occur as a result of pre- and postnatal chemical exposure, e.g. on reproduction, on the developing nervous system or on the developing immune system. This study evaluates endocrine endpoints in the juvenile and adult F1, which are not included in the 2-generation study (OECD TG 416) adopted in 2001.

****Project 4.64: Transcriptional Assay for the Detection of Estrogenic and Anti-Estrogenic Compounds using the MELN Cells.** Lead: European Commission. Project status and milestones: In vitro transactivation assay (2 protocols: one for a manual test method and one for a high throughput test method); Under validation; Expected to be integrated to TG 455 and TG 457 (BG1Luc ER TA).

Related to effects on biotic systems, that may also include endocrine disruption effects:

Project 2.1: EDTA Activity: New TG for Copepod Reproduction and Development. Lead: Sweden. Project status and milestones: Pre-validation completed, expert group meeting on 3-4 November 2005; Revised detailed ring-test plan circulated in November 2005; Ring-test between November 2005 and mid-2006; Submission of the draft report to the invertebrate expert group and the VMG-eco; Additional experimental work (plus report and statistical analysis) completed in 2007 Validation report (Phase 1) published in 2007; Expert group agreement on another ring test. Validation plan developed by Sweden and the US; Validation report (phase 2) endorsed at the 2011 WNT meeting; The US took the lead of an additional interlaboratory study; Outcome of additional work discussed at the VMG-eco meeting (November 2011) Comparison review between copepod and mysid; Experimental work expected to be completed in 2nd quarter 2012; Integrated Summary Report expected to be available end 3rd quarter 2012. Work ongoing.

Project 2.4: New TG 2-Generation Avian Toxicity. Lead: United States. Project status and milestones: Initial draft TG in 1999; decision to conduct an avian dosing study to be considered after DRP approval (Project 2.5 completed); Species comparison study circulated to the expert group in April 2005; Avian dosing study (July 2005) and revised draft TG (November 2005) presented at VMG-eco meeting in December 2005; Revised draft TG used in developing a protocol for a demonstration study; Report of the demonstration study expected before the end of 3rd quarter 2010; Revised Special Programme for Food Security (SPSF) expected to be submitted for the 2011 WNT meeting; Progress report at the VMG-eco in November 2011; Experimental work expected to be completed in May 2012; Post study/histopathology expected end 2012; Validation report expected in 2013.

Project 2.12: EDTA Activity: TG on a Medaka Life Cycle / Multi-generation Test (MMT). Lead: United States + Germany + Japan. Project status and milestones: Biateral work in Japan and the United States on medaka FLC/2-Gen in 2005-2006; Germany is working on a parallel project; 2007: preparation of a report comparing the various results available, based on 2 to 3 substances; Review of statistical issues circulated for comments in September 2007 to the fish drafting group and the VMG-eco; Revised review of statistical issues submitted to the VMG-eco meeting in January 2008; Development of a TG by Japan and the United States expected at a later stage; Revised SPSF submitted by the US and Japan agreed at the 2009 WNT meeting; Discussion of the approaches proposed and best course of action by the VMG-eco in December 2009 and by the fish expert group at a meeting in June 2010; Consensus Japan/USA 27 week-Medaka Multigeneration Test presented at the fish drafting group meeting held on 9-10 February 2011 in Japan; Progress report at the VMG-eco meeting in November 2011; Data

expected to be available for discussion by the Validation Management Group for Ecotoxicity Testing in December 2012 in Paris.

Project 2.13: EDTA Activity: New TG for Mysid Life Cycle Toxicity Test. Lead: United States. Project status and milestones: First proposal for a TG submitted in 2004; Pre-validation work completed in the U.S. in July 2005; Issue discussed at the 2nd Meeting of the invertebrate expert group, on 3-4 November 2005, and progress report at the VMG-eco in December 2005; Preliminary ring test results available. More validation needed; Optimization; US may pursue a national development of the test method depending on the interest of other member countries; Secretariat asked the WNT whether other countries than the US were interested in the project. If no country was interested, the project would be moved to Annex 1; Germany expressed interest in the project; Depending on the validation outcome, the SPSF will be revised and other partners will be invited to participate; Validation work completed (At that time there may be a need for a study comparing the copepod reproduction and development test (Project 2.1) with the mysid test; Comparison review between copepod and mysid; Outcome of additional work expected discussed at the VMG-eco meeting (November 2011); Experimental work expected to be completed in 2nd quarter 2012; Integrated Summary Report expected to be available end 3rd quarter 2012.

Project 2.31: EDTA Activity - TG on a Larval Amphibian Growth and Development Assay. Lead: United States + Japan. Project Status and milestones: Report at the VMG-eco meeting in November 2011; Validation expected to be completed in 2nd quarter 2012; Validation report expected to be available end 2012.

Project 2.36: EDTA Activity - New TG(s): Mollusc Reproductive Toxicity Tests – Development and Validation of Test Guidelines. Lead: Germany, United Kingdom, France, Denmark. Project Status and milestones: Annual meetings of lead countries; Pre-validation and Ring test under the oversight of the VMG-eco; Report on the pre-validation of Potymopyrgus and Lymnea tests at the VMG-eco in 2011 and2012; First draft TG expected to be ready for a WNT commenting round by the end of 2014.

Project 2.39: EDTA Activity – New TG: Xenopus Embryonic Thyroid Signalling Assay. Lead: France. Project status and milestones: Project approved by written procedure on 1 June 2011; Completion of a document summarising results from different demonstration studies in 2011; Project discussed at the VMG-eco meeting in November 2011; Comprehensive written validation plan, including a detailed protocol, distributed to VMG-eco and participating laboratories in June 2012; Inter-laboratory trial in 2012. Work ongoing.

****Project 2.7: New TG for Fish Embryo Toxicity test.** Lead: Germany. Project status and milestones: First draft of the TG and comprehensive background paper sent to the WNT for comments by 1 September 2006; Establishment of an expert group to address comments received; First teleconference of expert group on 29 January 2007 (decision on further validation steps), followed by two other conference calls; Meeting of the expert group on 9-11 October 2007 in Berlin; Expert group addressed the issue of performance; meeting arranged by ILSI/HESI at L'Oreal in Paris in March 2008 with a session on the OECD FET; Progress reported at WNT 20 meeting in 2008; Meeting of the expert group on 14-15 May 2008 in Berlin; Validation Management Group established to oversee the validation study that was initiated in May 2009. Validation report (phase 1) endorsed at the 2011 WNT meeting and published in 2011; Validation (phase 2) initiated in October 2010 with additional testing of 13 chemicals in 8 laboratories; Expert group meeting in Berlin on 16-17 February 2012; Phase 2 validation report will be submitted for endorsement at the WNT in April 2012; Phase 2 validation report endorsed by the WNT and submitted to the Joint Meeting for declassification before publication, 2nd draft TG sent to WNT for comments by 14 September 2012.

Adapted from OECD (2012) WORK PLAN FOR THE TEST GUIDELINES PROGRAMME (TGP) http://www.oecd.org/env/ehs/testing/August%20Work%20plan%20for%20the%20Test%20guidelines%20programme%20June%202013.pdf.
**= projects not mentioned by the OECD in the overview table 2 from document "Information on OECD Work related to Endocrine disruptors" (OECD, 2012f).

D. EPIDEMIOLOGICAL, FIELD AND (ECO)TOXICOLOGICAL INFORMATION

This annex describes types of studies that can be used to inform Level 1 of the OECD CF, OECD GD 150, (OECD, 2012a), and that may already be available in scientific literature prior to specifically conducting regulatory tests in subsequent levels of the CF.

D.1 Use of epidemiological information when characterising the hazards for humans of exposure to endocrine active substances

Because the main concern of EAS exposure and potential adverse effects relate to humans (e.g. as reviewed in the report on early development effects (WHO, 2012)), the conduct of epidemiological studies appears to be a logical and necessary complement to non-clinical studies to better characterise the nature and magnitude of the risk of endocrine disruption in humans.

However, general difficulties of epidemiological studies are related to their observational nature which is prone to a number of selection biases and methodological difficulties. More specifically, epidemiological studies aiming at deciphering a plausible causal relation between EAS exposure and various adverse effects should account for the following: 1) Environmental EASs are numerous and ubiquitous, 2) Humans are generally exposed to low levels and multiple substances and, 3) Other 'environmental' conditions such as lifestyle factors may come into play (Main et al., 2010; Martina et al., 2012).

Precise estimates of the exposure together with the identification of the critical developmental stage(s) to a particular EAS or combined exposure to multiple EASs are major challenges in studies aimed at assessing the effects of EASs in humans.

Designing relevant exposure measurements

A number of studies have focused on EAS perinatal exposure due to the possible irreversible impact of an exposure during this critical developmental period. In this particular case, the exposure of the individual at the time of disease manifestation is not important, but the exposure of the parents at time of conception and/or prenatal period is the one that needs to be assessed. Since accurate exposure data of the parents are hardly available it is consequently difficult to establish any causal association between early life exposures and adverse health effects in adulthood.

A unique measurement of exposure e.g. in urine, provides a snap-shot of exposure, but is not a good surrogate marker to estimate any long term exposure to ubiquitous substances that do not accumulate in the human body (e.g. phthalates). In addition, the total exposure may involve a repeated exposure to very low doses of different EASs.

Perhaps partly due to the above difficulties, several studies which have tried to quantify the exposure to various EASs have failed to find a link between exposure and adverse effects (e.g. Swan et al., 2005).

Health effects currently attributed to EAS exposure are often multifactorial, highly prevalent in the general population (e.g. obesity, diabetes, breast or prostate cancer, decreased fertility or congenital malformations), making the identification of an increased risk associated with the exposure to EASs only possible via the conduct of cohort studies with a very large sample size due to the need of controlling for a very wide range of possible confounding factors; e.g. diet, exercise, smoking status, social status (Vrijheid et al., 2012; Gehring et al., 2013) are examples of such studies. Significant changes in nutrient and food intake over the recent years also needs to be taken into account.

The problem of defining adequate study groups and to find controls in humans studies

It is often impossible to find negative controls, i.e. subjects who have not been or are not exposed to EASs and other substances. The absence of proven effects observed in humans following EAS exposure might be related to the low exposure levels but also to the difficulty to account for the multiple exposures from the environment.

Occupationally exposed workers represent an interesting group of subjects to assess the effects of some EASs in humans due to the predominant high levels of exposure to specific substances compared to environmental exposures. However, due to the ubiquitous nature of many of these substances finding control groups poses an important methodological problem. Inaccurate biomonitoring data could lead to misclassification (i.e. wrongly consider that subjects have not been exposed to EASs despite the fact that they have actually been), which in turn will induce a bias towards finding no effect between the active and the control groups. A lack of adequate biomonitoring data is a major limiting factor in preventing the identification of causal associations.

Independent of these difficulties, wide population surveys in combination with specific epidemiological studies will still be the most important approaches to identify adverse effects of endocrine disruption in the human population. Criteria for objectively evaluating the level of causality of associations observed in epidemiology have been formulated by Bradford Hill (1965) and include consistency, strength of association, dose response, time order, specificity, consistency of replication, predictive performance, biological plausibility (including evidence from non-clinical studies) and coherence.

Considering the uncertainties surrounding the effects of EASs on human health and limitations of extrapolation from non-clinical data, it can be concluded that the conduct of epidemiological studies, in spite of their inherent challenges, remains an essential component of the evaluation of possible human effects of EASs in large populations.

Advances in our understanding of critical windows of susceptibility (see section 4.7.2) and of the cumulative effects of multiple EASs acting in synergy (see section 4.7.3) add to existing concerns about the health impacts of EASs. The increased level of knowledge shows that the possible impacts of EASs on human health cannot be dismissed and should be taken into account in risk assessment (EEA, 2012).

D.2 Use of field information when characterising the hazards for wildlife of exposure to endocrine active substances

Laboratory studies have shown that a broad range of species from vertebrate and invertebrate taxa are susceptible to EASs; for many species, strong evidence exists to indicate that endocrine disruption is a widespread phenomenon in wildlife populations. Clear examples of male and female reproductive dysgenesis and of thyroid hormone disruption that can be linked, quite convincingly, to ED exposure have been reported in some wildlife (non-target) species. The causal relationship is, however, difficult to establish (EEA, 2012). For example, although an extensive number of studies have demonstrated endocrine disruption in wild freshwater and marine fish e.g. by Allen et al. (1999), the causality of the worldwide decline in amphibian populations and the possible effects and contribution of endocrine disruption to such decline remains uncertain (EEA, 2012).

In principle, there is no major difference between the evaluation of effects in the field of EDs and non-EDs. However, in view of the need to be confident that an adverse effect of an ED is likely to have consequences at the population level, the use of field data, if available, may be valuable. In the absence of such data, regulators must be confident of being able to extrapolate from laboratory data on

such endpoints as growth and reproduction to potential effects on populations, ideally but not necessarily through the use of population modelling. It is acknowledged that some effects, particularly those affecting individual behaviour, may not easily be apparent when observed at a population or subpopulation level.

There are two types of field information – field experiments, and monitoring data. The former approach involves the study of animal populations in treated artificial ponds (mesocosms), cages or field plots (e.g. Kidd et al., 2007). Through the use of replication and defined dosing with the substance of interest, this is essentially a field extension of laboratory experiments, and can provide reliable population-level data which firmly link the treatment with an adverse effect. In the case of EASs/EDs, it may also be desirable to measure biomarkers of exposure and/or mechanism in the exposed organisms (e.g. vitellogenin – VTG - induction in male fish as a response to oestrogens) so as to provide further evidence that the observed population-level impact is indeed associated with postulated endocrine activity. The main drawback of mesocosm, cage or plot experiments is their cost, and the complexity of evaluating the results. By contrast, effects on some aspects of animal behaviour, particularly those which are best expressed in the environment of a natural experimental arena, may be studied more effectively in the field than in the laboratory.

Field monitoring is probably the most powerful tool in establishing impacts at the population-level (e.g. population declines). Historically, field monitoring has been instrumental in illustrating the environmental hazards of chemicals like dichlorodiphenyltrichloroethane (DDT), Polychlorinated Biphenyls (PCBs) and mercury and their negative impacts on wildlife populations. Field monitoring is usually combined with chemical monitoring of the environment and the food chain. Correlation and multifactorial analysis are employed to assess the results. Such data are similar to human epidemiological information, and it may become necessary to employ the Bradford Hill criteria (Bradford Hill, 1965) to decide by WoE whether the substance of interest is likely to be responsible for the observed effects. Ultimately, this may involve returning to the laboratory or mesocosm to conduct replicated experiments to test hypotheses derived from the field evidence. For an example of field monitoring, see the case of tributytin antifouling paints (Matthiessen and Gibbs, 1998). It should be noted that while field studies can provide confidence that an ED is indeed able to cause population-level adverse impacts, there may be practical difficulties associated with this type of approach. In many cases, a judgement about whether test endpoints such as impaired reproduction are likely to cause population damage in the field will have to be made in the absence of field data or even of population modelling. In these circumstances, the expert judgement of the regulator or assessor will need to be deployed. A key question, for example, may be the degree of reproductive impairment (or other altered endpoint) which is likely to have consequences for populations. In other words, it may be insufficient merely to demonstrate a statistically significant effect on (say) fecundity – the effect must also be of sufficient magnitude so as to have biological relevance for the population (EFSA Scientific Committee, 2011).

D.3 Use of in vitro and animal studies when characterising the hazards for humans or wildlife of exposure to endocrine active substances

There is a wealth of *in vitro* studies based on cellular systems or isolated receptors. As information accumulates the strength of different *in vitro* assays in predicting endocrine disruption related health effects is becoming more and more evident. The US EPA ToxCast programme is using receptor assays and enzyme assays in a high throughput system (HTS) for screening chemicals for endocrine activity (U.S. EPA, 2012b). The OECD Guidance Document 150 (OECD, 2012a) lists a number of *in vitro* screens and tests (Level 2) to evaluate substances for potential endocrine disruption, discussing their applicability and limitations.

Animal studies have contributed considerably to our understanding of health disorders as a consequence of endocrine disruption in humans and wildlife. Reproduction-related disorders continue to dominate this area (for example, several EDs, either anti-androgenic or oestrogenic, can induce

various developmental anomalies of the male genital tract in experimental animals (Toppari, 2008), mirroring what has been observed in DES exposed humans) but neurodevelopmental and thyroid hormone–mediated disorders have also joined the list. The WHO report (WHO, 2012) on early development effects, lists more than 100 examples of endocrine disruption related effects in experimental animals caused by pharmaceuticals and industrial chemicals. Without doubt this constitutes a body of evidence that substances cause endocrine disruption resulting in adverse effects. Since the animal models so clearly show such effects, this is a strong indication that humans could also be at risk provided the exposure is adequate. These animal experiments were reported in general scientific literature (i.e. were not performed under a regulatory framework) and they may often be carried out with high exposure levels that are not necessarily relevant to humans. For regulatory purposes, whilst maintaining a full dose response curve, *in vivo* toxicological studies should be conducted at exposure levels relevant to human exposures.

The requirement for demonstrating adverse effects of an ED is manifest in the OECD conceptual framework for ED testing in the provision for higher Level *in vivo* tests (i.e. Levels 4 and 5) with the capacity to demonstrate effects on apical endpoints relevant to protection goals (e.g. population-relevant for ecotoxicological species). This obviously places a burden on animal testing in the near term, which cannot be avoided in the absence of robust methods for extrapolating data from *in silico*, *in vitro* or *ex vivo* methods to predict adverse effects *in vivo*.

In conclusion, experimental and mechanistic studies will continue to be the key sources of information when judging if a substance could act as an ED, and which by extrapolation could lead to adverse health effects in humans and wildlife. The fact that a substance in an *in vitro* assay is binding to an endocrine receptor, then interfering with the intracellular messenger system connecting receptor to target or resulting in an endocrine related response in a target cell must be taken as strong indication for endocrine activity. If a suitable animal model provides further indication for an endocrine-related adverse effect, this substance should be considered an ED.

GLOSSARY

Adverse effects	Change in the morphology, physiology, growth, development, reproduction, or life span of an organism, system, or (sub)population that results in an impairment of functional capacity, an impairment of the capacity to compensate for additional stress, or an increase in susceptibility to other influences.	IPCS
Androgens	Androgens are steroidhormones that help to develop sex organs in men. They also contribute to sexual function in men and women.	Endocrine Society
Apical endpoints	Results of an *in vivo* assay which describe a response by the organism as a whole, (e.g. fecundity or growth) which have possible implications for its biological fitness, rather than a response of the endocrine system alone (including physiological changes dependent on the endocrine system, such as Vitellogenin induction). Apical responses may or may not result from endocrine changes (e.g. fecundity may be affected both by some EDs and by some non-EDs).	OECD
Critical effect	The first adverse effect, or its known precursor, that occurs to the most sensitive species as the dose rate of an agent increases.	EPA Glossary
Dose response relationship	Relationship between the amount of an agent administered to, taken up by or absorbed by an organism system or (sub)population and the change developed in that organism, system or (sub)population in reaction to the agent.	IPCS
Expert system	A database of rules compiled by specialists (see section 4.2.2.6).	EFSA
Grouping of substances and read-across approach	Substances whose physicochemical, toxicological and ecotoxicological properties are likely to be similar or follow a regular pattern as a result of structural similarity may be considered as a group, or 'category' of substances. Application of the group concept requires that physicochemical properties, human health effects and environmental effects or environmental fate may be predicted from data for reference substance(s) within the group by interpolation to other substances in the group (read-across approach). This avoids the need to test every substance for every endpoint.	JRC/IHCP
Hazard	Inherent property of an agent or situation having the potential to cause adverse effects when an organism, system, or (sub)population is exposed to that agent.	IPCS
Hazard characterisation	The qualitative and, wherever possible, quantitative description of the inherent property of an agent or situation having the potential to cause adverse effects. This should, where possible, include a dose–response assessment and its attendant uncertainties. Hazard characterisation is the second stage in the process of hazard assessment and the second of four steps in risk assessment.	IPCS
Hazard identification	The identification of the type and nature of adverse effects that an agent has an inherent capacity to cause in an organism, system, or (sub)population. Hazard identification is the first stage in hazard assessment and the first of four steps in risk assessment.	IPCS

Hormone	Made by endocrine glands, hormones are chemical messengers that travel in the bloodstream to tissues or organs. They affect many processes, including growth, metabolism, sexual function, reproduction, and mood.	Endocrine Society
HTS	High-throughput screening (HTS) is a method for scientific experimentation especially used in drug discovery and relevant to the fields of biology and chemistry. Using robotics, data processing and control software, liquid handling devices, and sensitive detectors, HTS allows a researcher to quickly conduct millions of biochemical, genetic or pharmacological tests. Through this process one can rapidly identify active compounds, antibodies or genes which modulate a particular biomolecular pathway. The results of these experiments provide starting points for drug design and for understanding the interaction or role of a particular biochemical process in biology.	JRC/IHCP
In silico methods	The expression *in silico* is used to mean 'performed on computer or via computer simulation'. The phrase was coined in 1989 as an analogy to the Latin phrases *in vivo* and *in vitro* which are commonly used in biology and refer to experiments done in living organisms and outside of living organisms, respectively.	JRC/IHCP
Intact organism	Not *in vitro* systems, or castrated or ovariectomised test animals	EFSA
In vitro assay	Assay where whole live animals are not used. Systems used may include cell lines or subcellular preparations from untreated animals.	JRC/IHCP
In vivo assay	Assay where a whole live animal is treated. This may be a mammalian assay where individual animals are treated or a wildlife assay where a population of animals is treated.	JRC/IHCP
(Endocrine) Modality	A modality is an axis, pathway, signalling process or hormonal mechanism within the endocrine system.	EFSA
Mode of Action	A biologically plausible sequence of key events leading to an observed effect supported by robust experimental observations and mechanistic data. A mode of action describes key cytological and biochemical events – that is, those that are both measurable and necessary to the observed effect – in a logical framework.	IPCS
No observed adverse effect level (NOAEL)	Greatest concentration or amount of a substance, found by experiment or observation, that causes no adverse alteration of morphology, functional capacity, growth, development or lifespan of the target organism distinguishable from those observed in normal (control) organisms of the same species and strain under the same defined conditions of exposure.	IPCS
Oestrogens	Oestrogens are a group of steroid compounds that are the primary female sex hormones. They promote the development of female secondary sex characteristics and control aspects of regulating the menstrual cycle.	Endocrine Society
(Q)SAR	(Q)SARs are methods for estimating properties of a chemical from its molecular structure and have the potential to provide information on hazards of chemicals, while reducing time, monetary cost and animal testing currently needed.	JRC/IHCP
Read-across	'Read-across' is a technique of filling data gaps. To 'read-across' is to apply data from a tested chemical for a particular property or effect (e.g. cancer, reproductive toxicity) to a similar untested chemical. The read-across technique is often applied within groups of similar chemicals assembled for assessment using either analog approach (grouping based on a very limited number of chemicals) or category approach (grouping	JRC/IHCP

	based on a larger number of chemicals). In an analog/category approach, not every chemical needs to be tested for every endpoint. Source: EPA, Glossary of Terms, Methods of Toxicity Testing and Risk Assessment.	
Risk assessment	A process intended to calculate or estimate the risk to a given target organism, system, or (sub)population, including the identification of attendant uncertainties, following exposure to a particular agent, taking into account the inherent characteristics of the agent of concern as well as the characteristics of the specific target system. The risk assessment process includes four steps: hazard identification, hazard characterisation (related term: dose–response assessment), exposure assessment, and risk characterisation. It is the first component in a risk analysis process.	IPCS
Risk management	Decision-making process involving considerations of political, social, economic, and technical factors with relevant risk assessment information relating to a hazard so as to develop, analyse, and compare regulatory and non-regulatory options and to select and implement appropriate regulatory response to that hazard. Risk management comprises three elements: risk evaluation; emission and exposure control; and risk monitoring.	IPCS
Thyroid hormone	The thyroid gland makes T3 (triiodothyronine) and T4 (thyroxine), which together are considered thyroid hormone. T3 and T4 have identical effects on cells. Thyroid hormone affects heart rate, blood pressure, body temperature, and weight. T3 and T4 are stored as thyroglobulin, which can be converted back into T3 and T4.	National Cancer Institute dictionary
Validated Assay	A test method for which validation studies have been completed to determine the relevance (including accuracy) and reliability for a specific purpose. It is important to note that a validated test method may not have sufficient performance in terms of accuracy and reliability to be found acceptable for the proposed purpose (OECD, 2005a).	OECD
Weight-of-evidence	A process in which all of the evidence considered relevant to a decision is evaluated and weighted.	IPCS
Wild life	Non-target species. This term does not cover wildlife intended to be controlled by the application of regulated products (i.e. target species).	EFSA

ABBREVIATIONS

A	Apical
ADI	Acceptable Daily Intake
ADME	Absorption, Distribution, Metabolism, Excretion
ANSES	French Agency for Food, Environmental and Occupational Health & Safety
AOP	Adverse Outcome Pathway
AR	Androgen receptor
CF	Conceptual Framework
CMR	Carcinogenic, Mutagenic or toxic for Reproduction
DES	Diethylstilbestrol
DRP	Detailed Review Paper
EAS	Endocrine active substance
EATS	Oestrogen, androgen, thyroid or steroidogenic
EC	European Commission
ECETOC	European Centre for Ecotoxicology and Toxicology of Chemicals
ECHA	European Chemical Agency
ED	Endocrine disruptor
EDTA	Endocrine Disruptors Testing and Assessment
EDTA TF	Endocrine Disrupters Testing and Assessment Task Force
EEA	European Environment Agency
EMA	European Medicines Agency
ENVI	Committee on the Environment,

	Public Health and Food Safety (European Parliament)
EP	European Parliament
EQSD	Directive on Environmental Quality Standards
ER	Oestrogen receptor
EU	European Union
FFLC	Fish full life cycle
FP	Framework Programme
GD	Guidance document
HEAL	Health and Environment Alliance, NGO
HPA	Hypothalamus-pituitary-adrenocortical
HTS	High Throughput Screening
IPCS	International Programme on Chemical Safety
JRC	Joint Research Centre
M	Mechanistic
MCF-7	Michigan Cancer Foundation-7
MG	Multi-generation
MMGT	Medaka Multi-generation Test
NMDRC	Non-monotonic dose response curves
NOAEL	No Observed Adverse Effect Level
NOEC	No Observed Effect Concentration
NRC	National Research Council
NTP	US National Toxicology Program
OECD	Organisation for Economic Co-operation and Development
PAN Europe	Pesticide Action Network Europe, NGO
PBT	Persistent, Bioaccumulative and Toxic
PPAR	Peroxisome Proliferator-Activated Receptor
PPP	Plant Protection Products
(Q)SAR	(Quantitative) Structure Activity Relationship
RBA	Relative binding affinity
REACH	Regulation on Registration, Evaluation, Authorisation and Restriction of Chemicals
rtER	rainbow trout oestrogen receptor
SAAED	State of the Art Assessment of Endocrine Disrupters
SAR	Structure Activity Relationship
SCCS	Scientific Committee on Consumer Safety
SCENIHR	Scientific Committee on Emerging and Newly Identified Health Risks
SCHER	Scientific Committee on Health and Environmental Risks
SPSF	Special Programme for Food Security
STOT-RE	Specific Target Organ Toxicity-Repeated Exposure
STTA	Stably Transfected Transcriptional Activation
SVHC	Substances of Very High Concern
TDI	Tolerable Daily Intake
TG	test guidelines
TH	Thyroid hormone
ToR	Terms of reference
TR	Thyroid receptor
UNEP	United Nations Environment Programme
US EPA	The United States Environmental Protection Agency
VMG–Mamm	The OECD Validation Management Group for mammalian toxicity testing
VMG-eco	The OECD Validation Management Group for ecotoxicology testing
VMG-NA	The OECD Validation Management Group on non-animal testing
vPvB	Very persistent and very bioaccumulative
WFD	Water Framework Directive
WHO	World Health Organization
WNT	The OECD Working Group of National Co-ordinators of the Test Guidelines Programme
WoE	Weight-of-evidence