STATE OF THE ART ASSESSMENT OF ENDOCRINE DISRUPTERS

Final Report

Project Contract Number
070307/2009/550687/SER/D3

Authors: Andreas Kortenkamp, Olwenn Martin, Michael Faust, Richard Evans, Rebecca McKinlay, Frances Orton and Erika Rosivatz

23.12.2011

TABLE OF CONTENTS

0 EXECUTIVE SUMMARY

This report presents the results of a project “State of the Art Assessment of Endocrine Disrupters” which was commissioned through competitive tendering by the European Commission, DG Environment.

The report summarises advances in the state of the science since 2002 and maps out ways of dealing with endocrine disrupters in important pieces of EU chemicals regulation, such as e.g. the Plant Protection Product Regulation, PPPR (1107/2009), the new Biocide Regulation and the chemicals regulation, REACH (1907/2006).

During the last two decades evidence of increasing trends of many endocrine-related disorders in humans has strengthened. Although the correct description of disease time trends is often complicated by a lack of uniform diagnostic criteria, unfavourable disease trends have become apparent where these difficulties could be overcome. There are negative impacts on the ability to reproduce and develop properly. There is good evidence that wildlife populations have been affected, with sometimes widespread effects.

Multiple causes underlie these trends, and evidence is strengthening that chemical exposures are involved. Nevertheless, there are significant difficulties in identifying specific chemicals as contributing to risks. Especially where chemicals do not stay for long periods in tissues after exposures have occurred, it is impossible to detect associations when exposure measurements cannot cover periods of heightened sensitivity.

Extensive laboratory studies support the notion that chemical exposures contribute to endocrine disorders in humans and wildlife. Exposure during critical periods of development can cause irreversible and delayed effects that do not become evident until later in life. It is these toxicological properties that justify consideration of endocrine disrupting chemicals as substances of concern equivalent to carcinogens, mutagens and reproductive toxicants, as well as persistent, bioaccumulative and toxic chemicals.

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 endocrine disrupters 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. It is thus far not possible to infer the possibility of adverse effects from positive results in relatively cost-effective screening level assays.

For a wide range of endocrine disrupting effects, agreed and validated test methods do not exist. In many cases, even scientific research models that could be developed into tests are missing. 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 endocrine disrupters by EU Member States and other organisations revealed some commonalities and areas of agreement. Controversial are proposals to deal with endocrine disrupters on the basis of potency-based cut-off values derived from Regulation No 1272/2008 on classification, labelling and packaging of substances and mixtures (CLP). Such values are largely arbitrary and not scientifically justifiable.

Defining endocrine disrupters for regulatory purposes will have to rely on criteria for adversity and endocrine-related modes of action. Based on earlier proposals by various Member States and other organisations, including ECETOC, 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. Substances producing effects shown to be of no relevance for humans or wildlife can also leave the decision tree, but 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 approach. This weight-of-evidence 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. Regulatory decisions about endocrine disrupters will have to rely on weight-of-evidence procedures which are yet to be developed.

There are still enormous knowledge gaps that need to be addressed through research and development projects. Urgently needed are further methods for the identification of endocrine disrupters. Concerted efforts should be undertaken to identify the full spectrum of endocrine disrupters present in the environment and in human tissues.

The following recommendations are made:

- Implement recently updated or enhanced validated and internationally recognised test methods in the testing and information requirements for PPPR and REACH,
- Develop further guidance documents for the interpretation of test data,
- Consider the creation of a separate regulatory class “Endocrine Disrupter” (ED),
- Develop weight-of-evidence procedures that deal with the available evidence by weighing the criteria “adversity” and “mode of action” in parallel, but not by applying these criteria sequentially to exclude substances from the assessment,
- Consider potency, together with other criteria such as lead toxicity, specificity, severity and irreversibility in a weight-of-evidence approach. Abandon “potency” as a rigid and decisive cut-off criterion for endocrine disrupters of regulatory concern, for lack of prospect of reaching a consensus by purely scientific criteria,
- Create regulatory categories that stimulate the generation of the necessary data, including test methods that are not validated, beyond the OECD Conceptual Framework.

1 INTRODUCTION

Three pieces of European Community legislation deal explicitly with endocrine disrupters: The Plant Protection Product Regulation, PPPR (1107/2009); the chemicals regulation, REACH (1907/2006) and the new Biocidal Product Regulation, BPR (currently under negotiation).

The PPPR stipulates that active substances, safeners and synergists with endocrine disrupting properties that may cause adverse effects in humans cannot be approved for use unless the exposure of humans under realistic conditions of use is negligible. However, the PPPR does not detail how endocrine disrupters should be defined for the purposes of this regulation. The task of developing such criteria is the responsibility of the European Commission which is mandated to present draft criteria by 14 December 2013.

Under the REACH regulation, endocrine disrupters may be included under the authorisation scheme if they are deemed to be Substances of Very High Concern (SVHC) according to Article 57 (f). When the regulation was enacted in 2006, it was recognised that there was limited scientific knowledge about the effects of endocrine disrupters. Consequently, the European Commission is mandated with reviewing the provisions of REACH regarding endocrine disrupters (Art 138 (7)) by 1 June 2013.

The negotiation of the text of the Biocidal Product Regulation (BPR) between the Council and the Parliament has been finalised. It stipulates that substances considered as having endocrine disrupting properties that may cause adverse effects in humans or which are identified in accordance with Articles 57(f) and 59(1) of REACH shall not be approved, unless the risk to humans is negligible. The European Commission is required to specify the scientific criteria for the identification of biocides with endocrine disrupting properties no later than 13 December 2013.

With the aim of providing the scientific underpinnings for the process of defining criteria for endocrine disrupters, the present report, “State of the art assessment of endocrine disrupters”, was commissioned by the European Commission, DG Environment, through competitive tendering. The report describes the scientific state of the art in the field, as it developed during the last ten years, summarises the views of member state experts, interest groups and international organisations on endocrine disrupters and elaborates on options for dealing with chemicals with endocrine disrupting properties in relevant EU pieces of legislation.

1.1 TERMS OF REFERENCE, SCOPE OF THE REPORT

The overall objectives of this project were to analyse and summarise results of regulatory relevance of the scientific debate in the field of endocrine disrupting properties of substances, and to describe and characterise any relationships among the different levels of the expanded OECD conceptual framework. The specifications of the invitation to tender defined three tasks, and these form the *terms of reference* of this report:

Task 1: Analysing scientific literature on endocrine disrupters

The latest scientific literature was to be searched, and relevant results summarised, using as a point for departure the WHO report "Global Assessment of the State-of-the Science of Endocrine Disrupters" which was published in 2002. The knowledge gathered in the WHO document was to be included in the present report and was to be extended by findings not considered in the 2002 document.

Accordingly, literature searches were conducted, with additional information drawn from an analysis of EU projects, conference publications and opinions of relevant EU Scientific Committees.

The result of these activities is the **Summary of the State of the Science on Endocrine Disrupters** (Annex 1).

The "Summary" followed the structure of the WHO (2002) Global Assessment of Endocrine Disrupters, with its orientation on human health endpoint and effects in wildlife. The search for relevant literature targeted papers that appeared after publication of the WHO (2002) report. Because the effective literature cut-off date of the WHO assessment was the year 2000, papers with a publication date between 2000 and 2010 formed the basis of the Summary. Papers that appeared after December 2010 could generally not be considered, and older literature was referred to as appropriate. Further details about the scope of the Summary and the literature search strategy can be found in Annex 1.

The Summary of the State of the Science was completed in January 2011 and submitted to the European Commission services. In the subsequent stages of the project, the Summary was used to highlight results of regulatory significance (see below, Task 3). It was clear from the start that this would trigger re-examinations of the scientific literature, and the work process of summarising the state of the science leading up to the preparation of the present final report had to be iterative. Therefore, the content of the Summary, as completed in January 2011, was to be considered **preliminary** and a **"work in progress"** until finalisation of the project.

The Summary was reviewed by Commission experts from DG Environment and EU agencies. Their comments were received in April 2011. In the summer of 2011 the European Commission published the Summary, and as a result, the member companies of CEFIC, the umbrella organisation of European chemical manufacturers, and those of the European Crop Protection Association (ECPA) took the opportunity to publish combined comments on that Summary. These comments were published in September 2011. The UK Environment Ministry DEFRA also provided comments in September 2011.

The January 2011 Summary was expanded and modified over the summer of 2011. The result of this process is the final version of the Summary, presented as **Annex 1** to this report. It also incorporates alterations and additions made in the light of the comments from European Commission experts, CEFIC and ECPA, and the UK DEFRA.

Task 2: Analyzing approaches to assess endocrine disrupting properties of substances used in selected EU countries, in major competing economies of the EU and in international bodies

Approaches to dealing with endocrine disrupters in EU member states, major competing economies and international organisations were to be compiled and analysed. In discussions with European

Commission officials it was agreed to also consult experts from interest groups and non-governmental organisations.

The consultation of all the experts took the form of interviews which were conducted on the basis of a structured questionnaire designed to cover key topics relevant for regulatory approaches. The aims of the interviews were

- to ensure that the present State-of-the-art Assessment takes note of all relevant current, ongoing or planned activities and approaches in EU Member States regarding identification, testing, assessment and regulatory management of EDs, and
- to ensure that all points considered critical and important by Member State experts are properly reflected in the final report to the Commission.

The interviews followed an interview guide that was prepared in advance and was approved by DG ENV. The basic version of this guide was prepared for interviews with experts from EU Member States authorities. For interviews with other organisation this guide was adapted as appropriate. Altogether, 18 experts were consulted. The outcome of this task is presented as **Annex 2** to this report, where a description of further procedural details can also be found.

Task 3: Drawing conclusions and answering policy relevant questions

On the basis of the Summary of the State of the Science, an overview of scientific findings of regulatory relevance was to be prepared. Approaches to the assessment of endocrine disrupting properties were to be compiled and analysed comparatively, with an emphasis on summarising strong and weak points. The suitability and availability of tests for the identification of endocrine disrupting properties was to be analysed, in particular in terms of their ability to cover various endocrine disrupting mechanisms. Any links and relationships between the various levels of the OECD conceptual framework were to be described and analysed.

The core of the activities in Task 3 was a comparative analysis of the proposals of EU Member States for dealing with endocrine disrupters in EU Regulations.

The outcome of Task 3 is the present State of the Art Assessment of Endocrine Disrupters. Supplementary material to these efforts is presented in **Annex 3**.

1.2 STRUCTURE OF THE REPORT

The report is structured into seven main parts:

It begins with a consideration of definitions for endocrine disrupters, and their relevance in the EU regulatory context (Chapter 2).

Chapter 3 gives a description of the OECD conceptual framework for the testing of endocrine disrupters, and deals with tiered testing strategies that were developed in the USA and in Japan.

A summary of scientific findings of regulatory relevance to dealing with endocrine disrupters is presented in Chapter 4. It covers topics such as weight-of-evidence approaches, concepts for evaluating modes of action, the issue of low dose effects and thresholds, as well as critical windows

of susceptibility and irreversibility of effects. Endpoints and assays relevant for establishing endocrine disrupting properties are analysed, and an overview is given of chemicals of concern.

Chapter 5 summarises how the current European regulatory framework deals with endocrine disrupters. It focuses on three regulations, classification and labelling, REACH and PPPR.

Proposals developed by competent authorities of EU Member States for dealing with endocrine disrupters in the context of PPPR and REACH are described and analysed in Chapter 6.

Chapter 7 maps out options and recommendations for dealing with endocrine disrupters, particularly with respect to screening and testing strategies and in terms of decision criteria for defining endocrine disrupters in a regulatory sense.

Finally, the needs and requirements for research and development in the endocrine disrupter field are considered, with an emphasis on research that can directly inform the regulatory context (Chapter 8).

Chapter 9 are the three Annexes, with the Summary of the State of the Science on Endocrine Disruption in Annex 1, the summary of expert consultations in Annex 2 and supplementary material to findings of regulatory relevance in Annex 3.

2 DEFINITION OF ENDOCRINE DISRUPTING CHEMICALS

The WHO/IPCS definition[1] is referred to as a "working" definition in the "Community Strategy for endocrine disrupters"[2] and was generally acknowledged by the Member State experts consulted in connection with Task 2 of this project to provide a top-level definition that is applicable to both human health and ecotoxicological hazard and risk assessments (see Annex 2).

"*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.*"

Some important implications in terms of regulation of endocrine disrupters are left open to interpretation in this definition, and regulatory implications as highlighted in the "State-of-the-art summary of science on endocrine disrupters" (Annex 1) and by the consulted experts (Annex 2) are discussed further in this section. There are clearly two requirements for a substance to be defined as an endocrine disrupter, namely that of the demonstration of an **adverse effect** and of an **endocrine disruption mode-of-action**. Additionally, the definition implies **proof of causality** between the observed adverse effect and the endocrine disruption mode-of-action.

2.1 THE ENDOCRINE SYSTEM

The WHO/IPCS definition does not as such define the endocrine system and this has led to differential interpretation of the scope of endocrine disruption throughout the literature. The concept of endocrine disruption was first developed when it was observed that some environmental chemicals were able to mimic the action of the sex hormones oestrogens and androgens. It then evolved to encompass a range of mechanisms incorporating the many hormones secreted directly into the blood circulatory system by the glands of the endocrine system and their specific receptors, transport proteins and associated enzymes. Endocrine glands include the pituitary, thyroid and adrenal glands, and parts of the kidney, liver, heart and gonads and may signal to each other in series, thereby forming endocrine axes.

The three important endocrine axes are the hypothalamus-pituitary-gonad (HPG) axis, the hypothalamic-pituitary-adrenal (HPA) axis and the hypothalamic-pituitary-thyroid (HPT) axis. These axes describe the boundaries within which the endocrine system and endocrine disruption have been confined from the perspective of classical endocrinology. However, the scientific advances in our understanding of receptor signalling and molecular biology are continuously blurring the borders between the nervous system, immune system and endocrine system. The current scientific knowledge on receptor signalling was reviewed for the "State-of-the-art summary of science on endocrine disrupters" (section 3.1 Annex 1) and the following developments in the scientific

[1] International Programme on Chemical Safety. 2002. Global Assessment of the State-of-the-Science of Endocrine Disruptors. World Health Organization, Geneva, Switzerland.
[2] Commission of the European Communities. 1999. Community Strategy for Endocrine Disrupters. COM(1999)706. Brussels, Belgium.

understanding of signalling illustrate some of the points from which ambiguity over the definition of the endocrine system may arise;

- The same hormones or chemical messengers can be involved in "classical" endocrine signalling to more distant tissues as well as in local paracrine and autocrine regulation, or even in neurotransmission. An interesting example is that of acetylcholine. The role of acetylcholine as a neurotransmitter is well established. There is however also evidence that it acts as a non-neuronal signalling molecule in an autocrine or paracrine fashion and that it plays an intermediary role in the interactions of non-neuronal cells with endocrine hormones, growth factors, cytokines and also the neural system. Moreover, in certain plants, acetylcholine mediates the biological effects of light[1].
- 'Classical' hormones have been found to act not only via nuclear receptors but also membrane receptors comprising of G-protein coupled receptors whose ligands include catecholamines, prostaglandins, adrenocorticotropic hormone, glucagon, parathyroid hormone, thyroid-stimulating hormone, luteinising hormone; cytokine receptors whose ligands comprise tumour necrosis factor α, growth hormone, leptin; receptors with intrinsic enzymatic activity with the ligands insulin, epithelial growth factor, atrial natriuretic peptide, transforming growth factor β; and ligand regulated transporters whose ligands include acetylcholine. Such cell surface receptors are involved in rapid signalling and this is relevant to endocrine disruption as xenoestrogens for example have been shown to be able to disrupt the rapid effects of estradiol with different potencies to their effects on classical, genomic responses that regulate gene expression. It is now well established that rapid effects occur for every type of steroid hormone. The rapid effects of steroids are vulnerable to disruption by environmental chemicals but these effects are not typically measured in the assays that are available in the OECD testing framework (see 3. below).
- Receptor ligands can have diverse outcomes after receptor binding, including agonism, antagonism, acting as an inverse agonist, as partial agonist/antagonist or as a mixed agonist-antagonist, and as modulators. As well as the nature of the ligand, the outcome can be driven by tissue type and activation status.
- Nuclear receptors can be activated by second messenger signalling systems, instead of by binding a ligand agonist. Examples in the endocrine system include a potential role for ligand independent activation of the estrogen receptor in multiple cellular outcomes and in male-typical sexual differentiation of brain and behaviour and of the progesterone receptor in female sexual behaviour. Ligand independent activation provides a further opportunity for the integration of multiple signalling pathways and for chemical modulation: receptors that have functional actions when unliganded can have those actions perturbed by ligands simply through an alteration in their tonic, physiological role. The observation of ligand independence suggests a role for so-called 'orphan' receptors which may not possess a cognate ligand but instead may function as unliganded receptors. The arylhydrocarbon receptor for example is considered to have important physiological roles in the absence of a known ligand.

[1] Grando SA, Kawashima K, Kirkpatrick CJ, Wessler I. 2007. Recent progress in understanding the non-neuronal cholinergic system in humans. Life Sciences 80(24-25): 2181-2185.

An implicit understanding of the endocrine system or endocrine signalling can therefore span from the classical definition of the endocrine system to one that encompasses any type of receptor-mediated signalling. A further important question that arises in the ecotoxicological context is whether the term "endocrine system" should be interpreted in the very narrow sense of the hormonal system of vertebrates or whether it should include not only invertebrates, but also microbes or plants. This was highlighted by some of the interviewed Member State experts as having serious potential consequences in the context of pesticides and biocides regulation, as certain herbicides, so-called plant growth regulators, are designed to target weeds by disrupting plant signalling (Annex 2).

2.2 ADVERSITY

Adversity is an important concept of the WHO/IPCS definition and was introduced as a key criterion to differentiate between a mere endocrine modulator (elicits an adaptative reversible response in endocrine homeostasis) and an endocrine disrupter.

2.2.1 DEFINITION

As a consequence, the WHO/IPCS definition includes the term "adverse health effects", and from this stems the need to carefully define what "adversity" should mean in the context of endocrine disruption. WHO/IPCS has also defined "adversity"[1], and this adversity definition was applied to endocrine disrupters during a workshop organised by the German Federal Institute for Risk Assessment (BfR) in Berlin in November 2009[2] (Federal Institute for Risk Assessement, 2009), where it was further specified in the context of reproductive effects:

"*A change in morphology, physiology, growth,* reproduction*, development or lifespan of an organism which results in impairment of functional capacity or impairment of capacity to compensate for additional stress or increased susceptibility to the harmful effects of other environmental influences"*.

Efforts of subjecting chemicals with certain features to regulation cannot proceed without finding scientifically sound definitions of the effects in question (here: endocrine disruption), but how such definitions are to be applied in regulatory practice needs additional efforts. In particular, criteria are required for translating specific test outcomes into categories that reflect different degrees of strength of evidence etc. It is important to realise that such criteria do not directly derive from a definition of an endocrine disrupter, nor can definitions of adversity be translated directly into regulatory practice. For this reason, it makes sense to separate the task of defining an endocrine disrupter from efforts of implementing this definition in regulatory practice. An example for factors that lie outside the realm of (scientific) definitions of endocrine disrupters is the additional testing needs required for assessments of the "capacity to compensate for additional stress or increased susceptibility to the harmful effects of other environmental influences". These may conflict with the

[1] International Programme on Chemical Safety. 2004. IPCS Risk Assessment Terminology. World Health Organization, Geneva, Switzerland.

[2] Federal Institute for Risk Assessment. 2009. Establishment of assessment and decision criteria in human health risk assessment for substances with endocrine disrupting properties under the EU plant protection product regulation - Report of a Workshop hosted at the German Federal Institute for Risk Assessment (BfR) in Berlin, Germany, from Nov. 11th till Nov. 13th 2009. Berlin.

requirements for controlled experimental conditions and considerations of animal welfare in testing. Another question is whether adverse effects caused by endocrine disrupters require a different regulatory approach than adverse effects with other mode of actions.

Finally, a complication inherent in the WHO/IPCS definition for endocrine disrupters is that it is not descriptive (such as the definitions for carcinogens) but is strongly connected to mechanisms of action. Several of the Member State experts that were consulted pointed out that regulation must be based on assay outcomes, not mechanisms, and that adverse effects associated with endocrine disruption overlap with carcinogens and reproductive toxicants (Annex 2). With mutagens, however, there is a precedent of regulating on the basis of a mechanism. Evidence of mutagenicity is used to differentiate between genotoxic and non-genotoxic carcinogens and in the latter case leads to the assumption of no-threshold effects that precludes the derivation of Derived No Effect Level (DNEL)[1].

2.2.2 ASSAY REQUIREMENTS

The WHO/IPCS definition states that an adverse effect should be observed in an intact organism. In terms of regulatory requirements, this can only be interpreted to mean demonstration of a positive result in an *in vivo* assay. This will require repeated or chronic exposure regimens, and specifically exclude simpler screening test systems with castrated or ovariectomised animals.

The contention that classical toxicity test provide sufficient evidence of adverse effect was advanced by some experts and challenged by others on the basis that concern over endocrine disrupters results from delayed and irreversible effects on multiple target organs or tissue following exposure during critical windows of development (Annex 2). Evidence for such critical windows of exposure and whether relevant periods of development are included in the exposure regimen of various standard toxicity assays is discussed further in section 4.7.

Furthermore, as effects are sometimes only seen in the progeny of exposed animals, exposure regimens may additionally require tests to include two or more generations of animals. The associated financial burden may be substantial, and the ethical implications the additional numbers of animals that this requires may be serious, but these considerations have to be separated from the scientific requirements of testing, and are outside the scientific discourse.

2.2.3 ECOTOXICOLOGICAL EFFECTS

Although it is obvious that the WHO/IPCS definition was originally developed with consideration to human health effects, it is generally considered to adequately address the protection of ecological targets by extending the term “adverse health effect” to “(sub)populations” (Annex II). For legislation related to the protection or enhancement of biodiversity, an adverse effect needs to be demonstrated not in individual organisms but in a population. This raises the question of how the concept of adversity should be interpreted in an ecological context. In laboratory tests, effects on survival, growth and development, and reproduction in single species are generally regarded as ecologically relevant for the maintenance of wild populations. Relevant adverse effects may also

[1] European Chemicals Agency. 2007. Guidance for the preparation of an Annex XV dossier on the identification of substances of very high concern. ECHA. Helsinki, Finland.

include more subtle endpoints such as behaviour or increased susceptibility to naturally occurring stressors (Annex 2).

Another question that was raised by the consulted Member State experts was related to the interpretation of the term 'endocrine system' and whether this should be restricted to the vertebrate hormonal system or also include plants, microbes and invertebrate taxa. This has serious implications for the regulation of pesticides and biocides which are specifically designed to target endocrine processes in plants or pests and potential unacceptable risks in non-target organisms.

2.3 MODE OF ACTION

The WHO/IPCS definition leaves the interpretation of which modes-of-action should be considered to "**alter the function of the endocrine system**" relatively open. Nonetheless, evidence of modes-of-action other than estrogenicity/anti-estrogenicity, androgenicity/anti-androgenicity and thyroid disruption is acknowledged to be poorly addressed by the assays currently listed in the OECD Conceptual Framework. An OECD draft Detailed Review Paper on the State of Science on Novel In Vitro and In Vivo Screening and Testing Methods and Endpoints for Evaluating Endocrine Disruptors is discussed further in section 3.1.

Further, the definition does not explicitly address the issue of indirect endocrine toxicity, or when an effect on endocrine function is observed secondary to overt toxicity in other organs or systems. This is related to what is referred to as "**specificity"** or sometimes also "**lead toxicity**", to describe the requirement for endocrine disruption to occur at lower doses than other mechanisms of toxicity. The concept of lead toxicity is not without implications for the regulation of **mixtures**, which are explicitly included in the WHO/IPCS definition. While this was highlighted as a positive feature by some experts, one respondent argued that this was only relevant in the context of classification and labelling but not for substance-oriented legislation. The definition of mixtures in the CLP regulation is however limited to intentional mixtures and does not consider the unintentional mixtures characterising real environmental exposures. Groupings for the assessment of cumulative risks from simultaneous exposure to different substances are currently being discussed.

It is also important to note that a given substance may act via more than one mode-of-action and that the mode-of-action leading to an adverse effect may differ depending on the timing of exposure (critical windows of development), as well as the endocrine status of the organism (related to age and gender).

2.4 PROOF OF CAUSALITY

The use of the term "consequently causes" in the WHO/IPCS definition has been interpreted as a requirement for detailed information on the relationship between altered function of the endocrine system and the adverse effect. Concerns have been expressed that this phrase in the definition implies too high a level of proof, which will compromise its usefulness in the regulatory arena and might obstruct regulatory action due to restrictive use of its implied meaning[1]. Amendments to the

[1] Danish Centre on Endocrine Disrupters. 2011. Report on Criteria for Endocrine disrupters. Copenhagen, Denmark.

WHO/IPCS definition to allow for upstream events to be used instead of adverse effects were discussed during a Workshop on OECD countries activities regarding testing, assessment and management of endocrine disrupters[1]. No agreement was reached but some experts pointed out that this would allow for newer toxicological methods to be used. Some experts warned against any amendments at this stage to a consensual definition that has been discussed thoroughly at international level over many years. Others pointed out that it would be equally inappropriate to interpret the term "endocrine disrupting properties" as used in current EU legislation as equivalent to the WHO/IPCS definition.

It is highly unlikely that the information required by adverse outcomes pathway or key event analyses that would establish causality will be available for most suspected endocrine disrupters. In the case of endocrine disrupters this is compounded by the fact that disruption of endocrine processes may result in a complex pattern of effects. The fingerprint of effects of model compounds such as estrogens and androgens is however well characterised. Information on biological activity and/or upstream events could be used for categorisation of chemicals by read-across as suggested during a recent OECD workshop on using mechanistic information in forming chemical categories[2]. Such a move towards a better use of computational and high-throughput screening is mirrored in the United States[3] and can be related to ethical considerations regarding the number of animals necessary and prohibitive costs for the tests that would be required to fulfil such a high level of proof. With regards to the number of animals, concerns have also been expressed regarding the statistical power of standard test and their ability to detect relatively rare events such as reproductive tract malformation (section 3.2).

In this context it is relevant to consider that **REACH** (REGULATION (EC) No 1907/2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals) requires that a chemical with endocrine disrupter properties must cause **probable** serious effects before it can be designated a Substance of Very High Concern (SVHC), requiring authorisation. The **PPPR** (REGULATION (EC) No 1107/2009 concerning the placing of plant protection products on the market) demands sufficient evidence that a chemical **may cause adverse** effects before authorisation for use as a pesticide can be refused. It is a matter for interpretation whether this means that a less definite adverse effect than required in the WHO/IPCS definition may already fulfil the requirements laid down in REACH and PPPR, and whether, despite some uncertainty, a chemical may be nominated as an EDC SVHC under REACH or refused authorisation under PPPR.

In any case, it is important to realise that the WHO/IPCS definition does not take account of these legal subtleties, and that it was not designed to do so. Rather, the WHO/IPCS definition approaches the topic from a scientific viewpoint, with the very strict level of proof that is applied in science.

CHEMTrust. 2011. CHEM Trust's Contribution to the Ongoing Debate on Criteria for EDCs. United Kingdom.

[1] Organisation for Economic Co-operation and Development. 2009. Workshop Report on OECD Countries Activities regarding testing, assessment and management of endocrine disrupters. Advisory Group on Endocrine Disrupters Testing and Assessment (EDTA) of the Test Guidelines Programme (ed.). Copenhagen, Denmark.

[2] Organisation for Economic Cooperation and Development. 2011. Report of the Workshop on Using Mechanistic Information in Forming Chemical Categories. Series on Testing and Assessment No. 138. Paris, France.

[3] National Research Council. 2008. Science and Decisions: Advancing Risk Assessment. Washington D.C., United States.

These (scientific) criteria are not directly applicable to the regulatory arena, and additional reflections are needed to operationalise the meaning of “probable serious effects”.

2.5 DEFINITIONS OF POTENTIAL AND POSSIBLE ENDOCRINE DISRUPTERS

A definition for a potential endocrine disrupter was given alongside the WHO/IPCS definition.

“*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.*”

A recent development is the introduction of a definition for a possible endocrine disrupter at the OECD EDTA meeting in April 2011[1];

“A possible endocrine disrupter is 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”.

The definition for potential endocrine disrupters was sub-divided to account for large differences in the level of existing evidence for an endocrine mechanism, from *in silico/in vitro* data to results of *in vivo* screening assays where there is no evidence of adverse available from multigenerational assays.

In conclusion, the WHO/IPCS definition of endocrine disrupters is widely regarded as a useful basis for dealing with endocrine disrupters. The task of implementing this definition for regulatory purposes, however, should be separated from refining definitions of endocrine disrupters, and will require additional efforts.

2.6 ENDOCRINE MODULATION, ENDOCRINE ACTIVITY, ENDOCRINE MODULATORY ACTIVITY

It is recognised that many chemicals are capable of interacting with steroid receptors (“endocrine activity”), but whether this always leads to adverse effects is often unclear. The terms “endocrine modulation”, “endocrine activity” or “endocrine modulatory activity” are often used to make a distinction from “endocrine disruption” which is allied with adversity.

There is no universally agreed definition for these terms. “Endocrine modulation” is used to highlight a key feature of the endocrine system, namely that it can react to external challenges by feedback mechanisms that compensate for perturbances by chemical substances. In their draft opinion on the use of the Thresholds of Toxicological Concern (TTC) concept for endocrine “modulators”, the EFSA

[1] Organisation for Economic Cooperation and Development. 2011. Draft Summary Record of the Second Meeting of the Advisory Group on Endocrine Disrupter Testing and Assessment. OECD Paris, France.

Scientific Committee[1] emphasised the "…plasticity of endocrine homeostasis, characterised by a high level of compensatory feedback." "Endocrine exposures are handled by the body primarily by adaptive homeostatic mechanisms. Only if the body is unable to regulate exposures within its limits of homeostasis is the threshold of adversity crossed. In that case, adverse effects can occur, which is often referred to as endocrine disruption. Endocrine disruption-related toxicity may have specific features deserving special attention (e.g. high susceptibility of long-term developmental programming)."

The view that the body handles endocrine exposures primarily by adaptive homeostatic mechanisms is not necessarily applicable when exposures occur during windows of heightened susceptibility. Points where "the threshold of adversity" is crossed have not been defined in practice.

[1] EFSA Scientific Committee 2011, Draft Scientific Opinion on exploring options for providing preliminary advice about possible human health risks based on the concept of Thresholds of Toxicological Concern (TTC), Draft Scientific Opinion endorsed for public consultation.

3 FRAMEWORKS FOR REGULATORY TESTING AND SCREENING

3.1 THE OECD CONCEPTUAL FRAMEWORK

In 1996, the OECD established a Special Activity on Endocrine Disrupter Testing and Assessment (EDTA) at the request of the Member countries and the Business and Industry Advisory Committee to the OECD (BIAC) to ensure that testing and assessment approaches for endocrine disrupters would not substantially differ among countries.

The OECD conceptual framework was developed to support the testing and assessment of potential endocrine disrupters. It is intended to apply to both new and existing substances and different chemical sectors such as pharmaceuticals, industrial chemicals and pesticides. The framework was drawn up giving consideration to the views of Member countries that were gathered through a Questionnaire. Important in the development of the framework were the OECD's Appraisal of Test Methods for Sex Hormone Disrupting Chemicals[1], proposed testing schemes developed as part of relevant, national activities such as the then USEPA's Endocrine Disrupter's Screening and Testing Advisory Committee (EDSTAC), and research activities in Japan as well as industry initiatives such as those undertaken by the European Chemical Industry (CEFIC).

The initial framework has been revised by the EDTA Task Force at its meetings to reflect the OECD member countries' views. The latest version of the conceptual framework is given in Table 1. Its five levels of organisation are not anticipated to be used as a tiered testing strategy but as a 'toolbox' listing assays considered to provide different types of information regarding the hazards of a substance with regards to its potential endocrine disrupting properties. Level 1 is concerned with existing data and could be considered as a prioritisation stage that is beyond the scope of this report. A workshop on "OECD Countries Activities Regarding Testing, Assessment and Management of Endocrine Disruptors" held in Copenhagen on 22-24 September 2009[2] recommended that a guidance document on the assessment of chemicals for endocrine disruption should be developed by the EDTA AG. This was supported by the EDTA AG at its meeting on 17-18 May 2010. The draft guidance document[3] is now publicly available and its aims and organising principles are briefly reviewed in this section.

[1] Organisation for Economic Cooperation and Development. 2002. Appraisal of Test Methods for Sex Hormone Disrupting Chemicals. OECD Series on Testing and Assessment No. 21. Paris, France.
[2] Organisation for Economic Co-operation and Development. 2009. Workshop Report on OECD Countries Activities regarding testing, assessment and management of endocrine disrupters. EDTA. Copenhagen, Denmark.
[3] Organisation for Economic Cooperation and Development. 2011. Guidance Document on Standardised Test Guidelines for Evaluating Chemicals for Endocrine Disruption - Version 11. OECD. Paris, France.

Table 1. The OECD Conceptual Framework as published in the draft guidance documents. Please note that TGs 456, TG 234 and TG 443 have since been adopted.

	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 transcriptional activation (TG 455) • Androgen or thyroid transcriptional activation (If/when TGs are available) • Steroidogenesis in vitro (draft TG 456) • MCF-7 cell proliferation assays (ER ant/agonist) • Other assays as appropriate	
	Mammalian Toxicology	**Non-Mammalian Toxicology**
Level 3 *In vivo* assays providing data about selected endocrine mechanism(s) / pathway(s)[1]	• Uterotrophic assay (TG 440) • Hershberger assay (TG 441)	• Xenopus embryo thyroid signalling assay (When/if TG is available) • Amphibian metamorphosis assay (TG 231) • Fish Reproductive Screening Assay (TG 229) • Fish Screening Assay (TG 230) • Androgenized female stickleback screen (GD 140)
Level 4 *In vivo* assays providing data on adverse effects on endocrine relevant endpoints[2]	• Repeated dose 28-day study (TG 407) • Repeated dose 90-day study (TG 408) • 1-generation assay (TG 415) • Male pubertal assay (see GD 150 [*i.e.*this GD] Chapter C4.3)[3] • Female pubertal assay (see GD 150 [*i.e.*this GD] Chapter C4.4)[3] • Intact adult male endocrine screening assay (see GD 150 [*i.e.*this GD] Chapter Annex 2.5) • Prenatal developmental toxicity study (TG 414) • Chronic toxicity and carcinogenicity studies (TG 451-3) • Reproductive screening test (TG 421 if enhanced)	• Fish sexual development test (Draft TG 234) • Fish Reproduction Partial Lifecycle Test (when/If TG is Available) • Larval Amphibian Growth & Development Assay (when TG is available) • Avian Reproduction Assay (TG 206) • Mollusc Partial Lifecycle Assays (when TG is available)[4] • Chironomid Toxicity Test (TG 218-219)[4]

	• Combined 28 day/reproductive screening assay (TG 422 if enhanced) • Developmental neurotoxicity (TG 426)	
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]	• Extended one-generation reproductive Toxicity Study (draft TG 443) • 2-Generation assay (TG 416 most recent update)	• FLCTT (Fish LifeCycle Toxicity Test) (when TG is available) • Medaka Multigeneration Test (MMGT) (when TG is available) • Avian 2 generation reproductive toxicity assay (when TG is available) • Mysid Life Cycle Toxicity Test (when TG is available)[4] • Copepod Reproduction and Development Test (when TG is available)[4] • Sediment Water Chironomid Life Cycle Toxicity Test (TG 233) [4] • Mollusc Full Lifecycle Assays (when TG is available) [4] • Daphnia Reproduction Test (with male induction) (TG 211) [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 disrupters and some non-EDs. Those in Level 4 are partial lifecycle tests, while those in Level 5 are full- or multiple lifecycle tests.

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 based on a case by case basis, taking into account all available information, bearing in mind the function of the framework levels
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.

3.1.1 CONCEPTUAL FRAMEWORK GUIDANCE DOCUMENT

The guidance documents were drafted in order to support regulatory decisions related to the hazard characterisation of substances screened for endocrine disrupting properties. As the regulatory context for decisions varies across Member Countries as well as the use of chemicals being tested, the framework adopts a flexible approach to the interpretation of results. The guidance document describes several scenarios for a specific test outcome and available evidence. For each scenario, it recommends one further testing step, should this be deemed necessary to increase the evidence as to whether the substance is an endocrine disrupter.

Because of the countless possible scenarios, guidance was only developed for a limited selection of tests. The tests considered have either been validated for their potential to detect endocrine disrupting properties or are pending validation. The assays for which guidance has been drafted are shown in Annex 3. Further the guidance document covers the same endocrine modalities as those considered within the OECD conceptual framework, i.e. estrogen receptor mediated, androgen receptor mediated, thyroid hormone mediated and disruption of steroidogenesis.

The document recognises that effects on the hypothalamus-pituitary-adrenal axis, on the arylhydrocarbon receptor (AhR) pathway and on neuro-endocrine development **are not covered in the framework and will be missed**. Further, the *in vitro* mechanistic screens and *in vivo* screens and tests considered cover endpoints relevant for humans or vertebrate wildlife, specifically fish, amphibians and birds. Results from invertebrate test guidelines were not included due to the poor current understanding of endocrinology in most invertebrates, and the lack of screening endpoints specifically related to endocrine disruption.

Level 2: *In vitro* assays providing data about selected endocrine mechanism(s)/pathway(s)

The *in vitro* screening assays assembled in level 2 can provide qualitative information about a specific mode of action, and assays designed to detect binding to either the estrogen or androgen receptors have been validated, as well as some assays able to detect disruption of steroidogenesis. These assays tend to be biased towards false positive rather than false negatives, and this is intended to avoid the risks of false negatives. There is generally good concordance between *in vitro* and *in vivo* screening assays (level 3) for estrogenic and androgenic modes of action (see also below). A source of false negative (or false positive) results in *in vitro* systems is related to their lack of metabolic system. With the intention of widening the scope of these *in vitro* screens, a detailed review paper considered the inclusion of metabolising systems[1]. There are however some concerns about cytotoxicity and such metabolising systems are neither validated nor commonly applied. A possible solution would be to carry out *in vitro* metabolism prior to level 2 assays, however the relative activities of xenobiotic metabolising enzymes *in vitro* may still differ from their activities *in vivo* due to differential availability of cofactors, the stability of the enzymes or loss of subcellular compartments.

[1] Organisation for Economic Cooperation and Development . 2008. Detailed Review Paper on the Use of Metabolising Systems for *In vitro* Testing of Endocrine Disruptors. No. 97. OECD. Paris, France.

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

Assays at this level are intended to screen for the ability of a chemical to interact with the estrogen, androgen and thyroid hormone receptor mediated modalities in an *in vivo* context. They are also able to detect some non-receptor mediated modalities such as inhibition of iodination of thyroid hormones and steroidogenesis inhibition. In addition, these assays may capture the effects of metabolic conversions on the activity of test compounds, an aspect that is absent in level 2 *in vitro* assays.

To increase the sensitivity of these screens, they utilise castrated or ovariectomised animals in the case of the Hershberger and uterotrophic assay, respectively. This is to avoid that possible effects of the test chemical are obscured by the influence of endogenous hormones in the animal. The exposure regimens encompass only a short period of the animal's entire lifecycle and do not necessarily include exposure during critical windows of development that would expose the full spectrum of effects. As such, these assays are thought to be useful for revealing the capability of a chemical to interfere with the respective hormone receptors, but without necessarily indicating adverse effects. It is also argued that substances detected as actives in castrated or ovariectomised assays cannot be classed as endocrine disrupters, because they do not conform with a key requirement in the WHO/IPCS definition, namely that effects have to be demonstrated in intact animals.

The validity of this proviso is contended in cases where immature animals are used; their HPG axes are not yet capable of compensating for endocrine perturbations, and this can be interpreted as providing evidence of adverse effects following exposure during a critical window of development. The fish short term reproduction assay TG 229 includes apical endpoints which may be affected as a result of an endocrine or other mechanism of action. Conversely, a negative result in a level 3 assay does not exclude the possibility that the tested substance has endocrine disrupting properties through other endocrine mediated mechanisms.

Level 4: *In vivo* assays providing data on adverse effects on endocrine-relevant endpoints

Assays at this level provide information on numerous endpoints, with the possibility of detecting endocrine modalities not recognised in level 3 assays. Level 4 includes many assays originally not specifically designed to detect EDCs, such as repeated dose studies, nor are they validated to do so. An exception is the 28-day repeated dose toxicity test (TG 407), but the validation showed the assay to be quite insensitive to weak EDCs that act via the estrogen or androgen receptors. Although a positive result at this level is indicative of an adverse effect, the use of these assays is limited by the restricted number of endpoints measured and in some cases, by the lack of exposure during critical windows of development.

Level 5: *In vivo* assays providing data more comprehensive data on adverse effects on endocrine-relevant endpoints over more extensive parts of the life cycle of the organisms

At this level, assays provide data on adverse effects that may be due to endocrine disruption or other mechanisms, but the pattern of effects can be indicative of endocrine-mediated toxicity. Results of two-generation reproduction studies (TG 416) should nonetheless be interpreted with caution: some endocrine sensitive endpoints were added only in 2001 as a result of an update of the

technical guidance. Tests conducted prior to that date have a limited ability of detecting EDCs, simply owing to the absence of relevant endpoints. Further, some endpoints sensitive to endocrine disruption are not included even in the updated version of the two-generation reproduction study, such as nipple retention, anogenital distance at birth, and measurement of thyroid hormones. The new extended one-generation reproduction study includes those endpoints as well as neurodevelopment and immunotoxicity modules and the possibility to detect adverse effects not currently included in other validated tests. This test also requires that an increased number of pups be examined and its sensitivity is expected to be greater than the two-generation assay. Delayed effects that can manifest themselves with ageing such as premature reproductive senescence are currently not included in any guideline study. There is greater confidence that multigenerational ecotoxicity assays would be able to detect ED effects. **It is therefore recognised that even the latest guidance for level 5 assays has considerable gaps in covering endpoints relevant for the detection of endocrine disrupting chemicals**.

In order to test the application of the Conceptual Framework Guidance Document, three case studies were carried with prochloraz, perchlorate and 4-tert-octylphenol. These substances were selected because they had relatively large databases available and covered the same modalities as those addressed in the Conceptual Framework. Draft Case Studies for 4-Tert-Octylphenol[1] and Perchlorate[2] were made publicly available in October 2011. Both case studies find that recommendations given in the Guidance Document generally provide sound advice about data interpretation and possible next steps. It is interesting to note that even for well researched chemicals, no complete dataset is available. The case study of perchlorate in particular is argued to provide a good example of **the importance of assessing the weight of all available evidence.** These case studies also indicate that the lack of a comprehensive level 5 assay precludes definitive conclusions.

3.1.2 DETAILED REVIEW PAPER ON NOVEL TESTS AND ENDPOINTS

Another output of the Copenhagen Workshop[3] was a recommendation that a Detailed Review Paper be drafted to evaluate the effects of chemicals on other endocrine pathways and to review *in vitro* and *in vivo* test methods for additional signalling systems important for endocrine toxicity such as glucocorticoid receptors, AhR, peroxisome proliferator-activated receptors (PPARs), and other endocrine related nuclear receptors to be considered for incorporation into the conceptual framework. A draft document is now publicly available for comments until November 2011[4] and this

[1] Organisation for Economic Co-operation and Development. 2011. Guidance Document (GD) on Standardised Test Guidelines for Evaluating Chemicals for Endocrine Disruption (No. 150).Case Studies using example chemicals (4-Tert-Octylphenol). Draft v1. Document n° ENV/JM/TG/EDTA(2011)13. EDTA. Paris, France.

[2] Organisation for Economic Co-operation and Development. 2011. Guidance Document (GD) on Standardised Test Guidelines for Evaluating Chemicals for Endocrine Disruption (No. 150).Case Studies using example chemicals (Perchlorate). Draft v1. Document n° ENV/JM/TG/EDTA(2011)14. EDTA. Paris, France

[3] Organisation for Economic Co-operation and Development. 2009. Workshop Report on OECD Countries Activities regarding testing, assessment and management of endocrine disrupters. EDTA. Copenhagen, Denmark.

[4] RTI International. 2011. Draft Detailed Review Paper State of the Science on Novel In Vitro and In Vivo Screening and Testing Methods and Endpoints for Evaluating Endocrine Disruptors. OECD. Paris, France.

is briefly presented in this section. Further specifics of endpoints and assays covered in the Detailed Review Paper can be found in Annex 3.

The Detailed Review Paper describes assays that have been used to detect endocrine-disrupting effects of chemicals on more recently discovered estrogen, androgen and thyroid signalling pathways, e.g., signalling via membrane receptors, and neuroendocrine pathways. These latter pathways may function upstream to regulate the production of hormones that interact with nuclear receptors, or may act through the production of peptide hormones, which contribute directly to endocrine signalling. As mentioned in section 2.1, nuclear receptors in vertebrates include, amongst others, the corticosteroid receptors (e.g., mineralocorticoid, glucocorticoid), retinoic acid receptor (RAR), retinoid X receptor (RXR), vitamin D receptor (VDR), and PPAR. Some of their ligands such as vitamin D, retinoids or fatty acids do not fit the classical view of a hormone. Not all neuro-endocrine pathways are covered in the document. The focus of the document was informed by existing evidence of both susceptibility to disruption and assay procedures sufficiently developed for protocol standardisation and validation. The Detailed Review Paper is structured according to the pathways considered, namely; the hypothalamus-pituitary-adrenal axis, the somatotropic axis, the retinoid signalling pathway, the hypothalamus-pituitary-thyroid axis, the vitamin D signalling pathway, and the PPAR signalling pathway. Pathways are typically branched, rather than linear, with various intersections among the different pathways considered and this is perhaps best illustrated by Figure 1 reproduced from the draft detailed review paper[1].

Another section of the document is dedicated to the role of epigenetics in endocrine regulation and it highlights the important role of epigenetic processes in regulation of gene expression, with largely irreversible effects (see also Annex 1, 3.4).

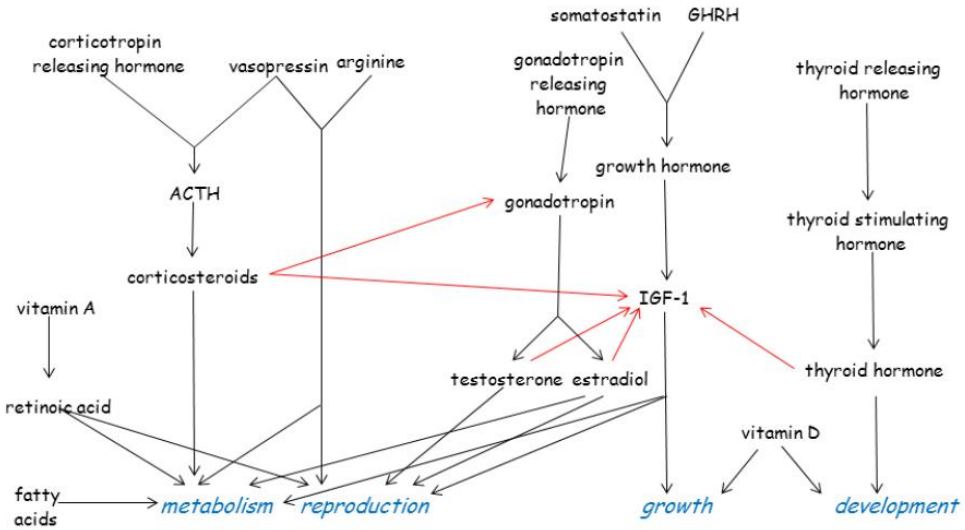

Figure 1. Neuro-Endocrine pathways known to be affected by EDCs resulting in symptoms of metabolic syndrome and disruptions in reproduction, growth, and development

[1] RTI International. 2011. Draft Detailed Review Paper State of the Science on Novel In Vitro and In Vivo Screening and Testing Methods and Endpoints for Evaluating Endocrine Disruptors. OECD. Paris, France.

3.1.3 CORRELATIONS BETWEEN TEST OUTCOMES AT LOWER LEVELS WITH THOSE AT HIGHER LEVELS

Dang et al. (2009) carried out a retrospective analysis of parameter sensitivity in multi-generation mammalian studies currently at level 5 of the OECD Conceptual Framework[1]. The publicly available literature on two-generation studies was searched for 161 category 1 and 2 reproductive toxicants on the ECB database and in PubMed. However, such information was available for only 18 substances and this illustrates that the absence of data is the main obstacle to meaningfully analyse any correlation between results of assays at level 2 of the OECD Conceptual Framework (*in vitro* assays) and level 5 assays. In contrast, a wider number of compounds have been screened *in vitro* for anti-/estrogenicity or anti-/androgenicity. The Endocrine Disrupter Knowledge Base and EU EDS database were searched for results of *in vitro* screens for the 18 compounds identified by Dang et al. Both *in vitro* data and results of a multi-generation study were available for 9 substances and a comparison of results in both types of assays is presented in Table 2.

[1] Dang Z-C, Rorije E, Hogen Esch T, Muller A, Hakkert BC, Piersma AH. 2009. Retrospective analysis of relative parameter sensitivity in multi-generation reproductive toxicity studies. Reproductive Toxicology. 28:196-202.

Table 2. Comparison of test outcomes at level 2 and 5 of the OECD Conceptual Framework

	Level 2	**Level 5 – Multi-generation assay**	
Substance name **Reproductive toxicity category**	*In vitro* assay type and result	LOAEL (mg/kg bw/day)	Critical reproductive effects
1,2,3-trichloropropane (Cat 1B - fertility)	**ER gene reporter**: Log(RP) = -1000	30	**P1**: average oestrus cycle length↑
Vinclozolin (Cat 1B – fertility and development)	**ER gene reporter**: Log(RP) = -10000 **ER Binding** Log(RBA) = -100 (mouse) Log(RBA) = -10000 (rat) **AR Binding** Log(RBA) = -2.5; -5 Ki > 700uM Metabolites M1: Ki = 92uM and M2: Ki = 9.7uM Vinclozolin, M1, M2: androgen agonist and antagonist effects in monkey kidney COS-1 and CV1 cells.	12	**F1**: day of preputial separation↑, ♂ abnormal nipple development; **P1**: prostate weight↓, prostate histology; **F2**: ♂ anogenital distance↓, ♂ abnormal nipple development
Potassium dichromate (Cat 1B- fertility and development)	**ER gene reporter**: Log(RP) = -10000	> 86	No reproductive toxicity LOAEL was determined in this study; **F1**: ♂ pup weight↓ 9-15%, ♀ pup weight↓ 11% at days 14 and 21 at the highest dose level, both not statistically significant.
DEHP (Cat 1B – fertility and development)	**ER gene reporter**: Log(RP) = -100; -10000 (both yeast) **ER binding**: Log(RBA) = -10000 **AR binding**: Log(RBA) = -5000	14	**P1**: small testis, testis aplasia, small epididymis, small seminal vesicles, seminal vesicles hypoplasia; **F2**: small testis, small epididymis
BBP (Cat 1B– development)	**ER Binding:** Log(RBA) = -1.8 (Human); -2.5 (mouse); -10000 (Rat) **E-screen:** Log(RPP) = 3.6 **Proliferation (ZR-75)**: 10 uM **ER gene reporter (Yeast)**: Log(RP) = -3.4; -4.2; -10000 1uM (MCF7) **AR Binding:** Log(RBA) = -2.1	250	**F1** and **F2**: ♂ anogenital distance↓
DBP (Cat 2- fertility Cat1B – development)	**ER Binding:** Log(RBA) = -2.58 (mouse); -10000 (rat) **Proliferation (ZR-75)**: 10 uM **Proliferation (MCF7)**: Log(RPP) = -4.08 **ER gene reporter**: ≥ 10uM (MCF 7) Log(RP) = -100; -10000 (yeast) **AR Binding:** Log(RBA) = -1.95 (rat)	♂ 52, ♀ 80	**P0**: dam weight during lactation↓; **F1**: live pups/litter↓; **F2**: live pup birth weight↓.

	Level 2	**Level 5 – Multi-generation assay**	
Substance name **Reproductive toxicity category**	*In vitro* assay type and result	LOAEL (mg/kg bw/day)	Critical reproductive effects
Nonylphenol (Cat2 – fertility and development)	**ER Binding:** Log(RBA) < -3.3 (ERα); -3 (ERβ); Log (RBA) = -0.5 (human); -0.5 (mouse); -1.05(human); -1.3 (human0; -1.5 (rat) **E-screen**: 10^{-13}M Log(RPP) = -2.5; -3.5 **ER reporter gene**: Log(RP) = -0.2 (human); -1.7 (yeast); -2.7 (yeast) **AR Binding:** Log(RBA) = -1.7 (rat)	150	**P0**, **P1**, **P2**, **F3**: ovarian weight↓.
BPA (Cat2 – fertility)	**ER Binding:** Log(RBA) = -0.5; -0.7; -1.25; -1.4; -1.3; -1.3; -2; -2; -2.1; -3; -3.3; -3.9; -4.2; -4.6 **E-screen:** Log(RPP) = -2; -2.8; -3.4; -4; -4.1 **ER reporter gene:** Log(RP) = -2; -2.2; -2.3 (yeast); -4 **PR induction (MCF7):** Log(RP) ≤ -4; = ; -3.3; -3.4 **AR Binding:** Log(RBA) = -2.4	50	**P0**: dam weight during gestation and nursing↓; **F1**: day of preputial separation↑; **P1**: dam weight during nursing↓; **F2**: day of preputial separation↑; **P2**: dam weight during gestation and nursing↓.
Acrylamide (Cat2 – fertility)	**ER reporter gene:** Log(RP) = -10000	5	**P0**: dam weight during gestation and nursing↓ 29%, implantations↓; **P1**: dam weight during gestation↓ 35%, implantations↓; **F1**: live pups/litter↓, ♂ pup body weight gain during lactation↓ 9%; **F2**: live pup body weight↓ 7%.

What is clear from this analysis is the dearth of data available to attempt any meaningful correlation between the lower and higher tier assays of the OECD Conceptual Framework. On the basis of this limited data, it is not possible to link the expression of a specific phenotypic expression of adverse effect to a specific endocrine mechanism. As testing requirements under REACH or PPPR are implemented, such an analysis may then become feasible.

3.2 TIERED TESTING STRATEGIES

The OECD Conceptual Framework is intended as testing framework (a "tool box"), and not a tiered testing strategy. Nevertheless, two OECD Member Countries have developed broadly similar tiered testing strategies.

Following the passage of the Food Quality Protection Act in the **United States**, the Environmental Protection Agency (USEPA) was required to develop a screening and testing program to determine human health effects of endocrine disrupting chemicals. The Endocrine Disruptor Screening and Testing Advisory Committee (EDSTAC) was established in 1996 to make recommendations on how to

develop the testing and screening program. EDSTAC published their final report in 1998[1] and recommendations were subsequently incorporated in the Endocrine Disruptor Screening Program (EDSP). The screening battery was designed to detect alterations in the developmental and reproductive processes controlled by the HPG and HPT axes. Based on a weight of evidence analysis of the results of Tier 1 assays, positive evidence in Tier 1 screening would lead to the substance tested to be considered a potential endocrine disrupter and subjected to further testing (Tier 2).

In Japan, The Advisory Committee on Health Effects of Endocrine Disruptive Chemicals has also developed a framework for testing of potential endocrine disrupters consisting of two tiers; namely, screening assays, including *in silico*, *in vitro* and *in vivo* assays, and definitive tests[2].

In both schemes, the goal of the first tier is to use screening assays sensitive enough to detect EDCs, whereas issues of dose-response, relevance of the route of exposure, sensitive life stages and adversity are resolved in the Tier 2 testing phase. When the USEPA first proposed a Tier 2 test for mammals, the multigeneration (2-generation) study was the only assay to have an appropriate exposure period covering the major developmental life stages of interest. This suggests that under such schemes the only assays that are appropriate for dose-response or potency assessment would be those listed in level 5 of the OECD conceptual framework. Notwithstanding the issue of exposure during critical windows of development, even multigenerational assays suffer from limitations imparted by their design: the limited number of offspring examined per litter in such multigenerational assay raises doubt over their statistical power to detect incidence of around 25% for endocrine-mediated effects such as reproductive tract malformations[3].

These considerations have important implications in terms of assessing the scope for detecting EDCs under the current testing requirements for REACH and PPPR (section 5).

[1] Endocrine Disruptor Screening and Testing Committee. 1998. Final Report, U.S. Environment Protection Agency.

[2] Organisation for Economic Co-operation and Development .2009. Workshop Report on OECD Countries Activities regarding testing, assessment and management of endocrine disrupters. Advisory Group on Endocrine Disrupters Testing and Assessment of the Test Guideline Programme. Copenhagen, Denmark.

[3] Hotchkiss AK, Rider CV, Blytone CR, Wilson VS, Hartig PC, Ankley GT, Foster PM, Gray CL and L. Earl Gray. 2008. Fifteen Years after "Wingspread"-Environmental Endocrine Disrupters and Human and Wildlife Health: Where We are Today and Where We Need to Go. Toxicological Sciences. 105(2): 235-259.

4 SCIENTIFIC RESULTS OF REGULATORY RELEVANCE

4.1 WEIGHT-OF-EVIDENCE APPROACHES

A general lack of consensus over the meaning of the term 'weight-of-evidence' (WoE) is widely recognised. The term is used in relation to the synthesis of evidence in the scientific peer-reviewed literature, as well as for classification and labelling purposes. The European Chemicals Agency (ECHA) has also published guidance on using weight of evidence within the context of the implementation of the REACH Regulation[1]. This guidance document, together with other WoE approaches which were specifically applied to the assessment of EDCs are discussed further in this section. WoE refers very generally to the synthesis or pooling of different lines of evidence. In the context of this report, it refers more specifically to the evidence of harm following exposure to a specific chemical substance. The term has been used to refer to a summary narrative of the result of hazard assessment where the methodological approach remains unspecified, systematic narrative reviews, criteria-based methods of causal inference, quantitative statistical techniques such as meta-analysis or as a label to a conceptual framework[2]. Conceptual frameworks for evaluating the evidence for a mode-of-action are addressed separately in 4.2.

Historically, WoE is distinct from an alternative approach referred to as 'strength of evidence' which analyses the degree of positive evidence from a subset of key studies that demonstrate a statistically significant result[2]. In contrast, WoE requires the synthesis of 'all' the evidence and to achieve this goal the analysis of evidence across several dimensions needs to be conducted, and this includes large or small, strong or weak, old and new studies over scales ranging from human populations to cellular systems. It particularly necessitates the combination of results from both human and animal studies. Traditionally, approaches use either epidemiological causal criteria or toxicological quality criteria and these are briefly described below as well as some aspects pertinent to the assessment of endocrine disrupters that have been discussed in the scientific literature.

4.1.1 EPIDEMIOLOGICAL CRITERIA OF CAUSAL INFERENCE

Epidemiological criteria of causal inference are almost invariably based on the so-called Bradford-Hill criteria or a modification thereof. In 1965, Austin Bradford-Hill, a British epidemiologist and statistician, published an article on the causes of occupational diseases featuring a list of nine 'considerations' for causation, given a body of statistically significant epidemiological evidence and some experimental toxicological evidence[3];

- Strength of the association - a quantitative measure of the relative risk estimate

[1] European Chemicals Agency. 2010. Practical guide 2: How to report weight of evidence. ECHA.Helsinki, Finland.
[2] Weed DL. 2005. Weight of evidence: A review of concept and methods. Risk Analysis 25(6): 1545-1557.
[3] Hill AB. 1965. The environment and disease: Association or causation? Journal of the Royal Society of Medicine, 58(5):295-300.

- Consistency of the observation across geographical, social or temporal scales
- Specificity of the association between exposure and disease
- Temporality – evidence that exposure precedes disease
- Biological gradient – evidence of response increasing with dose
- Plausibility of the association given what is known of the disease
- Coherence with the natural history and biology of the disease
- Experiment or recovery after exposure ceases
- Analogy – the consideration for the known effects of similar factors

Some of these aspects both in terms of the quality of epidemiological studies and aspects of causal inference have been argued to be coherent with a reductionist monocausal approach to disease aetiology and fail to adequately account for complex multicausal biological processes. Gee (2008) articulated the implications of these conflicting views of science for the application of WoE assessment in a regulatory context[1]. He argued that statistical adjustment for confounding factors may obscure associations by removing confounding factors that are in reality cocausal factors. With respect to the Bradford Hill criteria, he pointed to an asymmetry in their application to epidemiological studies: The presence of a specific criterion may provide robust evidence for an association, but the reverse is not true, absence does not provide robust evidence against causation. With respect to specific criteria, the following allowances can be made when considering the evidence to account for multiple causes that are pertinent to the assessment of evidence of effect following exposure to potential endocrine disrupters:

Temporality: an overall trend in a biological endpoint can be established under the influence of one cocausal factor before the emergence of other component causes.

Consistency: Inconsistency is to be expected as a result of the variability of complex biological and ecological systems.

Recovery: The persistence of some chemical substances and/or the latency or generational effects of developmental toxicants will delay potential recovery after cessation of exposure.

Biological gradient: The timing of exposure may be more important than the dose itself and evidence of a biological gradient may be further obscured by possible non-monotonic or low-dose effects.

Specificity: This criterion may not be relevant when considering multiple causes and multiple effects.

Strength of the association: In the context of multicausal diseases, a low relative risk can represent robust evidence of an effect if consistently replicated.

Further, Gee (2008) argued that greater weight should be given to the analogy criterion of the original Bradford Hill criteria and that such analogies could be used in conjunction with biological plausibility in a 'read-across' manner to fill some of the knowledge gaps in the toxicities of 'similar' substances. This should however be used with caution as some EDCs are known to affect multiple

[1] Gee D. 2008. Establishing evidence for early action: The prevention of reproductive and developmental harm. Basic & Clinical Pharmacology & Toxicology 102(2): 257-266.

receptor systems producing a complex profile of effects that is not necessarily identical to those interfering with a single hormone.

4.1.2 QUALITY CRITERIA FOR TOXICOLOGICAL STUDIES

The quality of toxicological studies is most commonly assessed using the approach developed by Klimisch et al. (1997)[1]. A scientific study is then assigned to one of four reliability categories:

1. **Reliable without restriction**, this generally applies to studies that conform to Good Laboratory Practice (GLP) or some other set of quality criteria.
2. **Reliable with restriction** applies to studies generally well documented and scientifically acceptable, but falling short of GLP in some measure.
3. **Not reliable**, methods used are either insufficiently documented or unacceptable.
4. **Not assignable**, when the documentation is insufficient for assessment (e.g. abstract only).

The OECD has described three terms used by Klimisch when referring to data quality as follows[2];

"***Reliability*** *- evaluating the inherent quality of a test report or publication relating to preferably standardised methodology and the way the experimental procedure and results are described to give evidence of the clarity and plausibility of the findings;*

Relevance *- covering the extent to which data and tests are appropriate for a particular hazard identification or risk characterisation; and*

Adequacy *- defining the usefulness of data for hazard/risk assessment purposes. When there is more than one study for each [Screening Information Data Sets] SIDS element, the greatest weight is attached to the study that is the most reliable and relevant. Robust study summaries are prepared for the highest quality or "key" studies.*"

Although more akin to a quality scoring scheme, this approach is still consistent with a WoE method using all the evidence, with some evidence weighted more reliable. It has been argued that the focus on GLP eliminates most modern scientific studies emerging from academic research, with the argument that scientific investigation per se entails non-standardised methods. On the other hand, the investigation of endpoints relevant to endocrine disruption requires specialist training and expertise that may not be readily available in all contract laboratories, even when they adhere to GLP.

Further, the scope for integration of human data in this scheme is limited and reflects a relative neglect of human data within the field of regulatory risk assessment of chemicals despite the greater weight imparted to human evidence for the purpose of classification of chemicals. This was

[1] Klimisch HJ, Andreae M, Tillmann U. 1997. A systematic approach for evaluating the quality of experimental toxicological and ecotoxicological data. Regulatory Toxicology and Pharmacology. 25:1-5.
[2] Organisation for Economic Co-operation and Development. 2007. Manual for investigation for High Production Volume Chemicals. Paris, France.

recognised by the European Centre for Ecotoxicology and Toxicology of Chemicals (ECETOC) who proposed a framework for the integration of human and animal data[1].

4.1.3 ECHA GUIDANCE ON WEIGHT OF EVIDENCE EVALUATION

Within Annex XI of the REACH legislation, use of a WoE approach is offered as an option to waive information/testing requirements of Annexes VII to X. In the ECHA guidance document on how to report weight of evidence, WoE is defined as "the process of considering the strengths and weaknesses of various pieces of information in reaching and supporting a conclusion concerning a property of the substance"[2].

Although reference is made to the Bradford Hill criteria for the evaluation of causality in epidemiological studies, the ranking of chemicals according to their endocrine potential[3] and evaluation of ecological risk[4], ECHA guidance only discusses Klimisch scores in some detail. Further guidance can be found in Chapter R.4 on evaluation of available information[5] that describes the adequacy requirements for human data, *in vitro* and *in silico* data. Emphasis is nonetheless on standard toxicology animal testing and identification of key studies, perhaps understandably as such guidance was not drafted for the purpose of assessing the weight of evidence for endocrine disrupting properties specifically.

In particular, examples of studies that are inadequate to qualify as key studies are given:

- "Problematic tests: Where a reasonable estimation of the exposure concentration cannot be determined then the test result should be considered with caution unless as part of a weight of evidence approach.
- Klimisch (2). 3 & 4 score studies
- Studies conducted according to non-standard guidelines"

Especially with respect to the last criterion, "studies conducted according to non-standard guidelines", the current scope and focus of WoE evaluations of chemicals under REACH is highly unlikely to allow the detection of endocrine disrupting properties.

[1] European Centre for Ecotoxicology and Toxicology of Chemicals. 2009. Framework for the Integration of Human and Animal Data in Chemical Risk Assessment. Technical Report No. 104. Brussels, Belgium.
[2] European Chemicals Agency. 2010. Practical guide 2: How to report weight of evidence. ECHA.Helsinki, Finland.
[3] Calabrese EJ, Baldwin LA, Kostecki PT and Potter TL. 1997. A toxicologically based weight-of-evidence methodology for the relative ranking of chemicals of endocrine disruption potential. Regulatory Toxicology and Pharmacology 26:36–40.
[4] Menzie CA, Henning MH, Cura J, Finkelstein K, Gentile J, Maughan J, Mitchell D, Petron S, Potocki B, Svirsky S, Tyler P. 1996. Special report of the Massachusetts weight-of-evidence workgroup: A weight-of-evidence approach for evaluating ecological risks. Human and Ecological Risk Assessment. 2(2):277–304.
[5] European Chemicals Agency. 2008. Guidance on information requirements and chemical safety assessment Chapter R.4: Evaluation of available information. ECHA.Helsinki, Finland.

4.1.4 EXAMPLES OF WOE APPROACHES FOR ENDOCRINE DISRUPTERS

The application of either epidemiological criteria for causal inference or quality criteria for toxicological studies to the evaluation of endocrine disrupting properties is best illustrated by briefly describing two examples, namely, the WoE framework applied to the evidence gathered in the State-of-the-Science on Endocrine Disrupters report published by WHO/IPCS in 2002[1], and the methodology used in the categorisation of substances on the candidate list of priority substances for evaluation of their role in endocrine disruption carried out by RPS BKH Consultants on behalf of DG Environment[2], each with a particular focus either on epidemiological criteria for causal inference or quality criteria for toxicological studies.

4.1.4.1 WHO/IPCS Global assessment of the state-of-the-science of endocrine disrupters

In chapter 7 of the WHO/IPCS Global assessment report, a framework based on modified Bradford Hill criteria was used for the qualitative assessment of relationships between exposures to potential EDCs and health outcomes. If it is recognised that a degree of scientific judgement is involved, the framework assesses two clearly stated hypotheses: associations between the outcome of concern (a specific health outcome or ecological species) and a putative stressor, and associations between exposure to the stressor and changes in endocrine-mediated events that may ultimately result in the outcome of concern.

The validity of the two hypotheses was then assessed according to five aspects of the scientific evidence:

- Temporality
- Strength of the association
- Consistency of the observations
- Biological plausibility of the effect
- Evidence of recovery following diminution of the stressor

The overall strength of evidence for the outcome, association with stressor and endocrine mode of action was then evaluated separately as weak, moderate or strong.

This type of framework allows the inclusion of relevant knowledge from a wide range of disciplines, integrating the current state of medical knowledge about a specific health outcome or subtle or indirect ecological effects. Because of its breadth of scope, such a holistic approach is less systematic. The method necessarily relies on expert judgement and may be criticised for its lack of transparency. It is nonetheless more appropriate for a global assessment of the evidence than approaches restricted to certain disciplines or types of studies.

[1] International Program on Chemical Safety. 2002. Global Assessment of the State-of-the-Science of Endocrine Disruptors. WHO, Geneva, Switzerland.
[2] Okkerman PC, van der Putte I. 2002. Endocrine disrupters: study on gathering information on 435 substances with insufficient data. DG Environment, Brussels, Belgium.

4.1.4.2 RPS-BKH Endocrine Disrupters database

It should be stressed that the RPS-BKH Endocrine Disrupters database is mentioned as an example of the application of toxicological WoE criteria and the Joint Research Centre - Institute for Health and Consumer Protection (JRC-IHCP) is currently developing the work already carried out towards a Web-based Information System on Endocrine Active Substances (EAS). Evidence of endocrine disrupting effects for 435 substances with insufficient data on the EU priority list was gathered both from regulatory sources and the peer-reviewed literature. Information about the systemic toxicity of the substances was also retrieved. The collected evidence was evaluated at two separate EU Expert meetings on endocrine disrupters held on 27-28 September 1999 and 9-10 September 2002.

The evaluation by experts consisted of the identification of key studies, the evaluation of the quality of key studies, and categorisation of the selected chemicals according to the evidence for their endocrine disrupting properties and the comparison of that information and ED potency and systemic toxicity data when available.

For the identification of key studies, experts could select up to three studies (if available) showing either positive or negative evidence on endocrine disruption. In cases where no key study could be identified, the substance was categorised as CAT3a or CAT3b.

The quality of the key studies was evaluated according to the following criteria:

Relevance of the effect parameter: with aspects such as relation ED effects with mechanistic cause

Test reliability:

- Use of validated protocols (analysis, test procedure);
- Experimental design: controls, concentration range;
- Test species: suitability, health, life stage;
- Analysis of results: statistics;
- Dose – Response relationship

Qualifying remarks: Coherence of the results of ED related tests

Additional considerations:

- Data availability (other ED tests);
- Comparison with systemic toxicity;
- ED potency:

The results of the evaluation distinguished 4 levels of data quality:

- DQ1: good data quality, fulfilling all (important) criteria;
- DQ2: sufficient data quality, study fulfilling most of the (important) criteria;
- DQ3: insufficient data quality, study cannot be used for identification;
- DQ4: not evaluated;

On basis of the identified key studies and their data quality, chemicals were categorised by RPS-BKH into four groups, according to the following definitions:

- Category 1. At least one study provides evidence of endocrine disruption in an intact organism. (Not a formal weight of evidence approach).
- Category 2. Potential for endocrine disruption. *In vitro* data indicate a potential for endocrine disruption in intact organisms. This also includes *in vivo* effects that may, or may not, be ED-mediated. May include structural analyses and metabolic considerations.
- Category 3a. No scientific basis for inclusion in list (ED studies available but no indications of ED effects)
- Category 3b. Substances with no or insufficient data gathered.

Experts had the option of adding "Qualifying remarks" referring to the coherence of the information in the database such as;

- 'other ED evidence is supporting';
- 'other evidence is lacking';
- 'other evidence is contradicting'.

Finally, considerations taking into account potency and systemic toxicity could be added with regards to the quantity of evidence, ED potency and comparison with systemic toxicity.

This approach focuses on toxicological evidence and as such is more consistent with approaches currently applied in the context of regulation of chemicals. It does not refer to Klimisch scores to assess the reliability of the data, but the criteria used are more or less equivalent, conferring greater weight to standardised protocols. Neither does it explicitly address the relevance and adequacy of the data, but such considerations are at least partially included in the suitability of the test species and probably the selection of key studies. The selection of key studies can be argued to be inconsistent with WoE methods and although the method is quite detailed it does also rely on expert judgement. Its main advantage is to offer some categorisation of EDCs on the basis of the available or selected evidence. The application of such 'bottom-up' approach does rarely consider the wider context of the human or ecological endpoints of concern.

4.1.5 IMPLICATIONS FOR THE REGULATION OF ENDOCRINE DISRUPTERS

The controversy about associations between specific chemicals and their purported endocrine disrupting properties can be directly related to the points highlighted throughout this section to divergent interpretations and quantitative or qualitative weights that were attached to different types of available evidence. The need to better integrate epidemiological evidence in the hazard characterisation of chemical substances has been recognised. The very existence of epidemiological evidence of effects linked to endocrine disruption in humans and wildlife is indicative of a failure of accurately predicting the toxicity of a given substance. WoE approaches to toxicological evidence ought to be better informed by the medical knowledge of the diseases they are supposed to predict and by a better understanding of comparative endocrinology and population/ecosystem dynamics.

Great emphasis has been placed on standardised validated methods and GLP practice when assessing the quality of toxicological studies. However, GLP and standardised protocols are no guarantee of quality if the most relevant sensitive endpoints or exposure during critical windows of development are not included. The systematic application of such quality criteria without any exercise of better judgement is likely to disregard critical evidence and hinder the implementation of timely preventative action. Moreover, it has been argued that reliability alone is not sufficient and that a set of criteria addressing the credibility of the work, regardless of the source of funding is equally necessary. Proposals for such a set of criteria were recently reviewed by Conrad and Becker (2011)[1].

Further, there is at present no universally accepted scheme for the classification of the results of WoE assessments. The International Agency for Research on Cancer (IARC) uses four categories to characterise the evidence for carcinogens; human, probable, possible and unlikely. Other approaches include those developed by the World Health Organisation in the field of air pollution and the United Nations Intergovernmental Panel on Climate Change (IPCC)[2]. The categorisation of endocrine disrupters requires the development of a similar scheme to classify evidence, as well as guidance on applying transparent and consistent concepts and terminology on cause/effect relationships.

4.2 CONCEPTUAL FRAMEWORKS FOR EVALUATING EVIDENCE OF MODE OF ACTION(S)

A specific application of WoE approaches is that of conceptual frameworks based on the evidence for the toxicological mode of action(s) of a given substance. Due to the requirement of establishing an ED mode of action that is enshrined in the IPCS/WHO definition, it is of interest to review available approaches together with their potential applications and limitations.

4.2.1 MECHANISM, MODE OF ACTION AND KEY EVENTS

The term “mode of action” is sometimes used interchangeably with the term mechanism of action. However in recent publications, a clear distinction between the two terms emerges. The mechanism of action is typically defined as the totality of mechanistic steps, whereas the mode of action refers to a less detailed sequence of key events within the mechanism. Key events can be defined as both measurable and necessary to elicit the apical toxicological effect of interest[3]. These definitions differ

[1] Conrad JW, Jr., Becker RA. 2011. Enhancing Credibility of Chemical Safety Studies: Emerging Consensus on Key Assessment Criteria. Environmental Health Perspectives 119(6): 757-764.

[2] World Health Organisation. 2005. Effects of Air Pollution on Children’s Health and Development. A Review of the Evidence. WHO, European Office, Bonn, Germany.
Intergovernmental Panel on Climate Change. 2007. Summary for policy makers. Climate Change: the Physical Science Basis. Contribution of Working Group 1 to the 4th Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom.

[3] Guyton KZ, Barone S, Brown RC, Euling SY, Jinot J, Makris S. 2008. Mode of action frameworks: A critical analysis. Journal of Toxicology and Environmental Health-Part B-Critical Reviews 11(1): 16-31.

in terms of their inclusion of toxicokinetic steps. These concepts are best illustrated by Figure 2, adapted from Guyton et al (2008)[1].

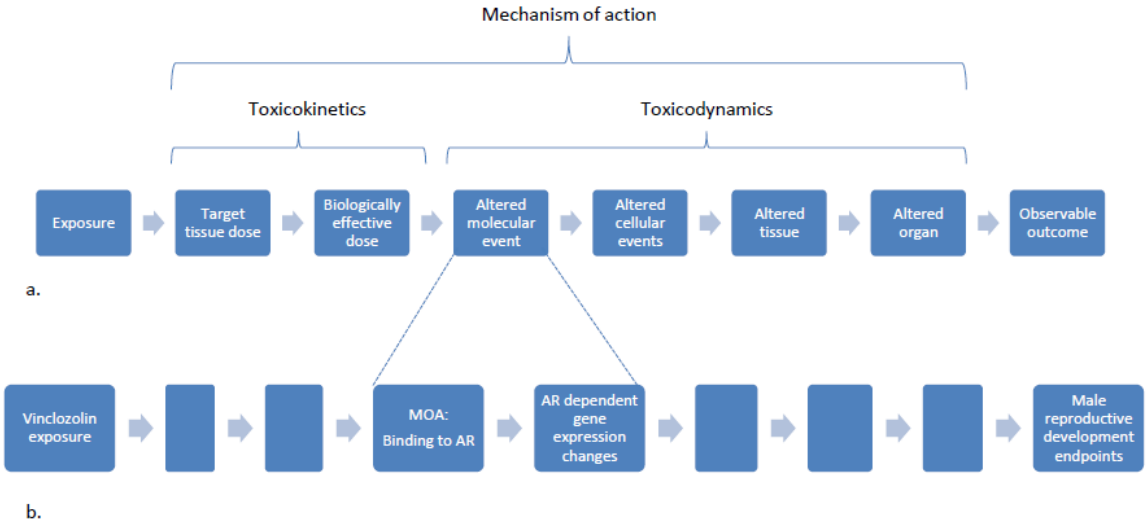

Figure 2. Illustration of mechanism and mode of action definitions. (a) Mechanism of action defined as all the steps between exposure and outcome. MOA is defined as "key event". (b) An example of the MOA is the pesticide chemical vinclozolin and its male reproductive developmental effects in the rat. The MOA/key event is binding to the androgen receptor (AR) (adapted from Guyton et al, 2008).

Information about the mode of action can inform decisions regarding the shape of the dose-response curve, particularly at low doses, the relevance of effects observed in experimental animals to humans, and the variability of the response within the human population including susceptible subgroups. An additional area of interest is the use of information on mode of action to inform cumulative risk assessment for mixtures.

4.2.2 MODE OF ACTION ANALYSIS METHODOLOGY

Frameworks for analysing data on the mode of action were first developed to assess the WoE for a carcinogenic mode of action in experimental animals and its relevance to humans[2]. Such a framework has recently been developed to assess noncancer modes of action[3] and is referred to in the Germany-United Kingdom proposal to assess human relevance for endocrine disrupting effects (see 6.4). The concept has been extended further to include quantitative descriptions of dose-response relationships for key events that may allow inferences about their shape, particularly at

[1] Guyton KZ, Barone S, Brown RC, Euling SY, Jinot J, Makris S. 2008. Mode of action frameworks: A critical analysis. Journal of Toxicology and Environmental Health-Part B-Critical Reviews 11(1): 16-31

[2] Boobis AR, Cohen SM, Dellarco V, McGregor D, Meek ME, Vickers C, et al. 2006. IPCS framework for analyzing the relevance of a cancer mode of action for humans. Critical Reviews in Toxicology 36(10): 781-792, Meek ME, Bucher JR, Cohen SM, Dellarco V, Hill RN, Lehman-McKeeman LD, et al. 2003. A framework for human relevance analysis of information on carcinogenic modes of action. Critical Reviews in Toxicology 33(6): 591-653, Sonich-Mullin C, Fielder R, Wiltse J, Baetcke K, Dempsey J, Fenner-Crisp R, et al. 2001. IPCS conceptual framework for evaluating a mode of action for chemical carcinogenesis. Regulatory Toxicology and Pharmacology 34(2): 146-152.

[3] Boobis AR, Doe JE, Heinrich-Hirsch B, Meek ME, Munn S, Ruchirawat M, et al. 2008. IPCS framework for analyzing the relevance of a noncancer mode of action for humans. Critical Reviews in Toxicology 38(2): 87-96.

low doses. These inferences are made on the basis of considerations of dose-dependent transitions in kinetic or dynamic processes related to the toxic response[1].

This WoE approach applies the Bradford Hill criteria to a postulated mode of action responsible for the toxicological effect of the test substance. In short, specific considerations include concordance between the dose-response relationships of key events and the toxicological response, temporal relationships between the key events, the strength, consistency and specificity of the association between the toxicological response and key events as well as the biological plausibility and coherence for the mode of action.

To assess the **human relevance** of the postulated mode of action, the evidence collated within the framework is queried in sequence for the sufficiency of evidence to establish a mode of action in the experimental animal, whether human relevance of this mode of action could be reasonably excluded on the basis of fundamental qualitative differences in key events between experimental animals and humans, and finally whether the relevance of the mode of action in experimental animals could be reasonably excluded on the basis of quantitative differences either in toxicokinetic or toxicodynamic factors between experimental animals and humans. The key default assumptions are considered to be protective as significant and convincing data are necessary to deviate from these default positions. However, these mode-of-action frameworks suffer from the same criticism as other WoE approaches in that the sufficient evidence is poorly if at all defined. Although they do offer some level of transparency about the rationale for the decision about human relevance, expert judgement is nonetheless still essential for that decision.

Interestingly, the approach has also been applied to **life-stage specific hazard characterisation**[2]. Relevant human and experimental animal studies are examined to identify critical windows of development for the observed outcome, lifestage-specific toxicokinetic and toxicodynamic data, mode of action information, variability and latency of effects. The aim is to determine the overall WoE of association between early life exposures and adverse outcomes on the basis of the adequacy, strength and completeness of the database. Because concerns following exposure to EDCs arise from the critical organisational role of hormones during development, the application of such an approach is particularly relevant to the hazard characterisation of EDCs (section 4.4). Further, a worked case study of the application of such a framework revealed that the consideration of multiple outcomes for a given mode of action in experimental animals uncovered that some endpoints are relevant to humans while others are not. This is important when considering those human diseases for which there is no adequate animal experimental model but evidence for a specific mode of action[3] (see also section 4.7).

While most approaches consider alternative modes of action for a given outcome, an often voiced criticism is that they do not allow for the integration of multiple modes of action. The underlying

[1] Guyton KZ, Barone S, Brown RC, Euling SY, Jinot J, Makris S. 2008. Mode of action frameworks: A critical analysis. Journal of Toxicology and Environmental Health-Part B-Critical Reviews 11(1): 16-31.
[2] U.S. EPA. 2006. A framework for assessing health risk of environmental exposure to children. Office of Research and Development, National Center for Environmental Assessment, Washington, DC; EPA/600/ R-05/093F.
[3] Seed J, Carney EW, Corley RA, Crofton KM, DeSesso JM, Foster PMD, et al. 2005. Overview: Using mode of action and life stage information to evaluate the human relevance of animal toxicity data. Critical Reviews in Toxicology 35(8-9): 663-672.

assumption appears to be that modes of action are mutually exclusive and the possibility that different modes of action may act in an interactive manner is not considered in currently available approaches. Another weakness is the lack of consideration of the known causes of human disease and the potential for chemical exposures to act additively with background exposures, disease or endogenous processes.

Finally, it should be recognised that the meaningful application of such frameworks to the hazard characterisation of chemicals will require extensive data that is unlikely to be available for most chemicals. It would therefore require a substantial effort to expand the knowledge base. It nonetheless offers an interesting approach to assess evidence from a viewpoint that specifically addresses hazardous properties that underlie the level of concern over EDCs such as effects at low doses and exposure during critical windows of development. The possibility of adapting such a framework to assess interspecies differences could also be applied in order to integrate evidence of effects from both human toxicology and ecotoxicology.

4.3 LOW DOSE EFFECTS AND THRESHOLDS

In connection with efforts to characterise the risks associated with EDCs it has been argued that the current risk assessment paradigm needs modification or has become obsolete, because EDCs elicit effects at doses much lower than normally used in regulatory testing. The “low dose” hypothesis of endocrine disruption[1] has expressed two separate, although connected, aspects of the issue:

(1) Risks to human health are feared at current exposure levels, because several studies have described effects of EDC in animal experiments at doses that approach the exposures experienced by humans.

(2) Non-monotonic dose response relationships have been observed with EDCs for certain endpoints. It has been argued that this challenges an assumption implicit in current risk assessment, namely that effects seen at high doses can be used for extrapolations into the low dose range. With non-monotonic dose response curves, this key assumption is no longer tenable, and it has been proposed that “low dose” testing should be performed routinely, in order to provide the basis for more protective points of departure (e.g. NOAELs or benchmark doses) that are subsequently used for deriving human reference doses.

The “low dose” phenomenon has been studied with a limited number of EDCs that exhibit varying modes of action (Annex I, 3.2). The validity of many of the observations has been the topic of intense debates, mainly due to issues relating to their reproducibility (see Annex I, 3.2 and the debates in[2].) These disputes are likely to continue in the foreseeable future, with prospects of a resolution uncertain, and dependent on the provision of further experimental evidence.

[1] vom Saal FS and Hughes C. 2005. An extensive new literature concerning low-dose effects of bisphenol A shows the need for a new risk assessment. Environ Health Perspect 113:926-933.

[2] Witorsch RJ. 2002. Low-dose in utero effects of xenoestrogens in mice and their relevance to humans: an analytical review of the literature. Food Chem Toxicol 40: 905; Kamrin MA. 2007. The “low dose” hypothesis: Validity and implications for human risk. Int J Toxicol 26: 13-23; Hayes TB. 2004. There is no denying this: Defusing the confusion about atrazine. BioScience 54: 1138-1149

Quite independent of the EDC "low dose" discourse, a wider-ranging discussion of estimating dose-response relationships in the low dose range has taken place outside the field of endocrine disrupter research which nevertheless has important implications for the regulation of EDCs[1]. This wider discussion has assumed urgency with recent epidemiological studies of very large populations (up to several 100,000), where thresholds were **not** observed, independent of whether cancer or non-cancer outcomes were analysed. Instead, risks increased linearly with dose in the low dose range. These observations have been made in studies investigating the effects of ozone, tobacco smoke, nitric oxide and sulphur dioxide, particulate matter and lead. On the basis of these epidemiological findings, experts increasingly question the current dichotomy in risk assessment, which deals with non-carcinogens by assuming thresholds, and maintains that the risks associated with carcinogens decrease with dose, but in a threshold-independent fashion. It has been proposed to put this dichotomy aside and to deal with pollutants in a uniform manner where threshold-independent action can also be assumed for non-carcinogens.

Two aspects of importance for the regulation of EDCs have emerged from this wider debate: (1) Attempts to categorically define modes of action that can inform low dose extrapolations, and (2) consideration of conditions that render thresholds not applicable for human populations.

4.3.1 CATEGORIES OF MODE OF ACTION THAT CAN INFORM LOW DOSE EXTRAPOLATIONS

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. Additional complicating factors are related to normal biological variation and the limited power that is available with the size of dose groups normally used in toxicity testing[2]. For these reasons, criteria for assuming threshold-independence of certain classes of pollutants derive from considerations of their mode of action, as for example with genotoxic carcinogens, where it is assumed that small numbers of irreversible events (mutations) can form the starting point of malignancies.

However, a thorough understanding of the events that lead from exposure to disease, and the modes of action involved in these steps is lacking for most environmental pollutants in general, and for EDCs in particular. A comprehensive elucidation of all the mechanisms involved is so data-extensive, that application to risk assessment remains quite a distant prospect. Conversely, the current concepts of mode of action in risk assessment are too general to be useful in informing low dose extrapolations.

These difficulties have given impetus to proposing categories of modes of action that are more amenable to low dose extrapolations, by considering the reversibility of key events, the rates of repair, if there is reversibility, and by examining the irreversibility of key steps. Three generic

[1] NRC 2009. Science and decisions. Advancing risk assessment. The National Academies Press. Washington DC. ISBN – 1-: 0-309-12046-2; White et al. 2009. State-of-the-science workshop report: Issues and responses in low dose extrapolation for environmental health risk assessment. Environ Health Perspect 117: 283-287.

[2] Slob W. 1999. Thresholds in Toxicology and Risk Assessment. International Journal of Toxicology 18:259-268; Scholze M and Kortenkamp A. 2007. Statistical power considerations show the endocrine disrupter low dose issue in a new light. Environ Health Perspect 115 Suppl 1: 84-90.

categories have been suggested[1]: (a) low dose reversible (e.g. irritants), (b) low dose irreversible (e.g. mutagens) and (c) chronic cumulative, irreversible events (e.g. neuronal loss in Parkinson's disease).

In applying such categories to EDCs, it becomes imperative to examine whether there is evidence for mode of action categories "low dose irreversible" or "chronic cumulative, irreversible". These issues are considered below (section 4.4).

4.3.2 ABSENCE OF THRESHOLDS AT THE HUMAN POPULATION LEVEL AND IMPORTANCE OF PRE-EXISTING EXPOSURES

Although the events that lead to exposure-related diseases may be non-linear in the low dose range, thresholds are obscured when the analysis is conducted at the human population level. Even under the assumption of thresholds for individuals (this will forever remain hypothetical, because thresholds cannot be verified at the individual level, even if present) thresholds are obscured at the population level by inter-individual variations in sensitivity and by background exposures or endogenous exposures that also have an impact on the health endpoint in question[2]. As a result, dose-response relationships often appear low dose linear at the population level, without evidence for a threshold, even though thresholds may exist for individuals. Population dose-response relationships then reflect a multitude of individuals' thresholds, with the consequence that a threshold cannot be established for the population (see the excellent discussion of this topic in Slob (1999)[3]).

A point of immediate relevance for EDCs concerns background exposures and endogenous exposures that play a role in disease processes. This scenario applies to pollutants that mimic the action of endogenous hormones, for example estrogens. Because of pre-existing internal exposures to steroidal estrogens, it can be inferred that any quantum of externally added estrogenic agent adds to the internal load, thereby exhibiting activity in a threshold-independent fashion. This is an important consideration for the role of estrogens in breast cancer (see Annex 1, 5.1), during the programming of the neuroendocrine system and timing of puberty (Annex 1, 4.2) and the role of xenoestrogens in influencing the sex determination in turtle eggs[4].

4.4 CRITICAL WINDOWS OF SENSITIVITY AND IRREVERSIBLE EFFECTS

Hormones are key factors for the proper development of a multiplicity of organ systems and tissues, with those of the reproductive tract, the brain and the neuroendocrine system being the most prominent. Accordingly, there is ample evidence for the exquisite sensitivity of the developing organism to chemical exposures that can interfere with normal hormone action during critical stages

[1] White et al. 2009, ibid
[2] NRC 2009, ibid
[3] Slob 1999, ibid
[4] Sheehan DM et al. 1999. No threshold dose for estradiol induced sex reversal of turtle embryos: how little is too much? Environ Health Perspect 107: 155-159.

of development. In many cases, the impact of the interfering chemicals is irreversible and stays with the affected organism for the rest of its life. There is also often a considerable delay between the time of exposure and the point when effects become manifest.

Some prominent examples include:

- The action of chemicals capable of interfering with androgen action during the male programming window in fetal life; this includes androgen receptor antagonists such as certain dicarboximide, imidazole and azole pesticides, and certain phthalates. Some read-outs of diminished androgen action in experimental animals only become apparent in adult life; this includes malformations of reproductive organs. The effects are largely irreversible (Annex 1, 4.1).
- Epidemiological studies show that exposures to dioxin (TCDD) in perinatal life have a negative impact on semen quality. When exposure occurred during puberty, the opposite effect manifests itself, while exposure during adulthood has no influence on semen quality[1].
- Estradiol and estrogenic chemicals can interfere with the KiSS peptide system in rodents in neonatal life, with influences on the timing of puberty (Annex 1, 4.2).
- The development of the female reproductive system is programmed *in utero* and can be disrupted at this stage by undue signalling from chemicals such as DES, with multiple and irreversible consequences (Annex 1, 4.5.3).
- Many hormonal cancers, including breast, prostate, testis, ovarian and endometrial cancer are thought to have their origins in fetal life and pubertal life. During these life stages, there is heightened sensitivity to chemical exposures implicated in these cancers (Annex 1, 5).
- The action of thyroid hormones during development in the womb is essential for many developmental landmarks, including the development of the brain and the neuro-endocrine system. Disruption of thyroid action by chemical exposures at this stage of development can have detrimental and irreversible effects (Annex 1, 6.1).
- There are many examples of vulnerable life stages in wildlife species, including lobsters, amphibians and reptiles that are exquisitely sensitive to the influences of EDCs (Annex 1, 7).

4.4.1 IMPLICATIONS FOR TESTING AND REGULATION

If the effects of EDCs are to be identified, it is essential that exposures cover windows of increased susceptibility in assays that incorporate the relevant endpoints. Many assays currently in use for the identification of carcinogens and reproductive and developmental toxicants do not incorporate the endpoints most sensitive to endocrine disruption, nor is administration of the test chemical at the relevant periods of sensitivity prescribed in all cases (see also 4.7 below). The events leading to the discovery of anti-androgenic effects of phthalates are instructive in this respect:

The standard testing regimens for reproductive and developmental toxicity used by the US NTP at the time did not involve administration of test compounds during the period of the male programming window. Furthermore, demonstrations of anti-androgenic effects require that all males in a litter are examined, but the NTP standard teratology protocol necessitated the testing of

[1] Mocarelli P, Gerthoux PM, Patterson DG, Milani S, Limonta G, Bertona M, Signorini S, Tramacere P, Colombo L, Crespi C, Brambilla P, Sarto C, Carreri V, Sampson EJ, Turner WE, Needham LL. 2008. Dioxin exposure, from infancy through puberty, produces endocrine disruption and affects human semen quality. Environmental Health Perspectives 116:70-77.

only one pup per litter. As a direct consequence of these conditions, the anti-androgenic effects of phthalates were long overlooked and only came to light by "accident"[1].

One motivation for including persistence and bioaccumulative properties of chemicals in the criteria that qualify for categorisation as substance of very high concern (SVHC) in REACH Art 57 was an aspect of irreversibility: Once emitted into the environment, the presence of such substances in environmental media cannot be reversed. If hazardous properties come to light at a later stage, harm to human health and wildlife cannot be prevented.

An analogous consideration can be used to justify the inclusion of EDCs in Art 57 as substances of concern similar to those listed in Art 57 (a-e): Because of their propensity to interfere with development and reproduction during key life stages, in often irreversible ways, the concern is that populations cannot be protected when the damage is diagnosed long after the causative exposures have taken place.

4.5 MIXTURES

There is good evidence that several EDCs can work together to produce combined effects. Especially when exposure is to multiple chemicals simultaneously that are capable of affecting the same endpoint, combination effects can occur at doses where each chemical individually is without detectable effects. From a regulatory point of view, it is therefore of great importance to have information about the spectrum of EDCs that are present in relevant exposure scenarios. This information is currently fragmentary, and this lack of information makes it likely that the full extent of risks associated with EDCs might be underestimated.

4.6 THRESHOLDS OF TOXICOLOGICAL CONCERN

The threshold of toxicological concern (TTC) approach is a **screening** tool that was developed in order to assess substances of unknown toxicity present at low levels in the diet on the basis of knowledge of the chemical structure of the substance concerned and information on human exposure. It uses generic human exposure threshold values (also called TTC values) that were established for substances grouped according to their chemical structure and likelihood of toxicity. The classification scheme for human exposure threshold values most widely used is that described by Cramer et al. (1978) based on the metabolic and toxicological information available at that time.

The EFSA Scientific Committee has evaluated the relevance and reliability of the TTC approach, and recently published a draft opinion for public consultation[2]. It found that the approach is applicable to substances for which the chemical structure is known but there are few or no relevant toxicity data. It recognises that revision and refinement of the classification scheme would be timely. It also

[1] NRC, 2008. Phthalates Cumulative Risk Assessment – The Tasks Ahead. Committee on Phthalates Health Risks, National Research Council, National Academy of Sciences, Board on Environmental Science and Technology, National Academy Press, Washington, DC

[2] EFSA Scientific Committee. 2011. DRAFT Scientific Opinion on Exploring options for providing preliminary advice about possible human health risks based on the concept of 5 Threshold of Toxicological Concern (TTC). European Food Standard Agency. Parma, Italy.

concluded that many endocrine mediated adverse effects involving reproduction, development and thyroid function would be adequately covered by the existing TTC values. The report also stressed that the TTC approach would not normally be applied when there is a legislative requirement for submission of toxicity data, and it is therefore of limited relevance in the context of this report.

4.7 ENDPOINTS, MODES-OF-ACTION AND ASSAYS

Pertinent findings published after the release of the WHO/IPCS Global Assessment of the State-of-the-Science on Endocrine Disrupters in 2002 have been summarised, with the aim of preparing the ground for an analysis of results of regulatory relevance to the scientific debate about endocrine disrupting properties of chemical substances. This summary followed the structure of the WHO (2002) Global Assessment of Endocrine Disrupters, with its orientation on human health endpoint and effects in wildlife. The Summary of the State of the Science was completed in January 2011 and amended in the light of comments received after its publication. This amended "Summary" can be found in Annex 1 of this report.

In order to evaluate the extent to which endpoints, modes of action and assays that have been examined in the scientific literature can be addressed by regulatory approaches, the analysis of scientific results was organised according to specific assay endpoints. For this purpose, the relevant information included in three documents was collated, namely the scientific summary found in Annex 1, the OECD Conceptual Framework Guidance Document[1] and the Detailed Review Paper on Novel Endpoints and Assays[2]. The assay endpoints mentioned in the OECD Conceptual Framework or in the Detailed Review Paper were matched to the human health and wildlife endpoints considered in the scientific summary in order to gain insights on the comprehensiveness of the Conceptual Framework and related guidance. This process was informed by other information gathered in the scientific summary. Evidence for one or more modes of action, for critical windows of susceptibility as well as any information on the human relevance of experimental models for human health endpoints or the relevance of an effect or mode of action for population effects for wildlife endpoints were examined. The results are presented in the tables that can be found in Annex 3 of this report. The most salient points that inform different states of knowledge and the competence of regulatory testing methods are recapitulated in the following section following the human health and wildlife endpoint structure adopted in the scientific summary.

[1] Organisation for Economic Cooperation and Development. 2011. Guidance Document on Standardised Test Guidelines for Evaluating Chemicals for Endocrine Disruption - Version 11. OECD. Paris, France.
[2] Organisation for Economic Cooperation and Development . 2008. Detailed Review Paper on the Use of Metabolising Systems for *In vitro* Testing of Endocrine Disruptors. No. 97. OECD. Paris, France.

4.7.1 HUMAN HEALTH

4.7.1.1 Reproductive health

4.7.1.1.1 Male Reproductive Health

Concerns regarding reported declines in male reproductive health, particularly semen quality, and the proposition for a testicular dysgenesis syndrome were instrumental in bringing endocrine disruption to the fore of the regulatory agenda. This was followed by a significant research effort and similarly significant scientific advances have been made in the field. There is good coherent evidence that suppression of androgen action by interfering with the uptake of steroid hormone precursors into foetal Leydig cells, blocking the androgen receptor, or by inhibition of steroid-converting enzymes, during the fetal male programming window, can interfere with male reproductive development (see section 4.1 in Annex 1 and tables 1 and 2 in Annex 3). There are differences in steroidogenesis between rats and humans, but in general the processes underlying male development are remarkably similar in both species and the rat is generally seen as an appropriate model for possible effects in humans. Further, in the rat model, there are sensitive and specific endpoints available such as anogenital distance and nipple retention. In the field of endocrine disruption, and perhaps regulatory toxicology more generally, this is a rather unique situation as the rest of this section will illustrate. Nevertheless, the measurement of said endpoints is required only in the extended one-generation reproduction study (TG 443) that has recently been adopted. From a regulatory perspective, the only obstacles to the inclusion of these endpoints in testing requirements are consideration for animal welfare and the cost of multigenerational mammalian assays.

There is also some recent evidence for other mechanisms that may lead to androgen suppression in fetal life, particularly inhibition of PGD2 synthesis and its subsequent role in differentiation of Sertoli cells, or a role for retinoids and IGF1 in spermatogenesis.

4.7.1.1.2 Female Reproductive Health

In contrast, if some of the epidemiological and toxicological experimental data suggest that endocrine disrupters are involved in female reproductive disorders, the demonstration of endocrine disrupting mechanism has remained elusive (see sections 4.2-4.7 in Annex 1 and tables 4, 5, 7, 8, 10, 11, 13, 14, 16 and 17 in Annex 3). It is yet unclear, for example with regards to precocious puberty or fecundity, whether the potential effects of endocrine disrupters result from disruption of central regulation of the female reproductive cycle or local effects on the ovary or mammary gland. For certain disorders such as polycystic ovaries syndrome, medical and toxicological evidence suggests a potential influence of excess androgens on fetal programming. There is also increasing evidence that adverse pregnancy outcomes could be related to oxidative damage as a result of disruption of the endocrine processes in implantation. Epigenetic regulation of Hox genes have also been implicated in endometriosis and uterine fibroids. The involvement of endocrine processes in the disease outcomes is often derived from medical knowledge of those disorders. The number of chemical messengers and receptors such as, not only estrogens and androgens, but also progesterone, glucocorticoids, prostaglandins, retinoids, the AhR or kisspeptin involved offer as many potential pathways for disruption. Whether the redundancy of signalling pathways makes the

female reproductive system more susceptible or more resilient to chemical injury is at this stage purely speculative. There is either indirect medical evidence or direct experimental toxicological evidence that disease outcomes can be influenced during critical period of development, not only *in utero* but also during the prepubertal or pubertal period. However convincing evidence of defined critical periods of development is restricted by the limited current medical understanding of the aetiology of those diseases. This is further compounded by, or arguably a result of, the lack of adequate model for some of those diseases due to critical differences between rodents and humans, particularly with reference to negative feedback on the hypothalamus by ovarian steroids in juvenile rodents but also because they are oestrous animals and do not menstruate. There are nonetheless consistent associations between exposure to endocrinally active chemicals and specific endpoints in rodents that have been exploited for the development of assays such as an advancement of age at vaginal opening or an increase in uterine weight following exposure to estrogenic compounds. Other female reproductive endpoints such as histopathology or gross morphology of reproductive organs included in assays are generally not necessarily specific to endocrine disruption. In the extended one-generation reproduction toxicity study (TG 443), endpoints related to adverse pregnancy outcomes are only measured in the parent generation (and the second generation, if triggered) and there is therefore a possibility that any effect on gestation resulting from exposure during a critical period of development rather than exposure during gestation may be missed. Effects on reproductive senescence are not monitored in any validated assays. A number of *in vitro* assays have been recently developed to detect the effect of chemical exposures on the ovary or placenta, and they will detect the effects of endocrine disrupters but will not allow assertion regarding the mechanism of toxicity.

In summary, the current state of knowledge would in all likelihood prevent the establishment of a clear causal link between an endocrine modality detected in an *in vitro* assay and an adverse effect in an *in vivo* even multigenerational reproduction assay for most endpoints.

4.7.1.2 Hormonal cancers

Issues surrounding the use of adequate animal models of hormonally sensitive cancers also mar the detection of hormonally-mediated carcinogenesis. The influence of hormonal factors, chemical exposures and advances in the medical understanding of hormonally mediated carcinogenesis support the plausibility of a role for exposure to endocrine disrupters in breast, prostate, testicular and to some extent thyroid cancers. The modalities of chief interest with respect to hormonally-mediated carcinogenesis are estrogenic, androgenic and exposure to thyroid stimulating hormone in the case of thyroid cancer. There is also convincing evidence that morphogenesis during fetal life is a period particularly sensitive to disruption at least for the testis and prostate. For the breast, puberty and potentially nursing may also be sensitive. There are at present no suitable animal models of prostate, testis or thyroid cancers (see section 5 in Annex 1 and tables 19 and 20 in Annex 3). The ACI rat is the only rat strain that develops mammary tumours with high incidence when exposed to DES, steroidal estrogens, including estradiol or equine estrogens used in HRT. The ACI strain is not highly susceptible to other types of neoplasms, and therefore not routinely used in validated carcinogenesis assays. The Noble rat is a good model for studying hormone-induced prostate cancers, but metastases are rare in this strain. Surrogate markers such as histopathological changes in the mammary gland, Leydig cell nodules/hyperplasia or prostate dysplasia are measured in repeated dose toxicity studies (TG 407) and multigenerational reproduction toxicity studies (TG 416

and TG 443) and those are included in the draft guidance document for the OECD Conceptual Framework. However, only in multigenerational studies will the animals be exposed during critical windows of susceptibility. Further measurements of thyroid hormones and thyroid stimulating hormone are only carried out in the extended one-generation reproduction study (TG 443).

4.7.1.3 Metabolism and development

Testing for neurotoxicity is not new in itself, but the interest in the susceptibility of the neuroendocrine system to chemical exposures during development is relatively recent. This may be related to the fact that for a long time, the brain and fetus were both thought to be relatively protected from chemical exposures by the so-called blood-brain barrier and placental barrier. These notions are increasingly being challenged. As a result, if a battery of functional behavioural endpoints are included in repeated dose toxicity studies (TG 407 and TG 408), these are unable to detect effects from exposure during brain development. The interest in neuroendocrine regulation of the immune system is also very recent. Accordingly, developmental neurotoxicity and immunotoxicity modules have been added to the extended one-generation reproduction study (TG 443). Although guidance on the interpretation of test results is included in the OECD Conceptual Framework guidance documents, it does not extend to the endpoints included as part of these two modules. Earlier interest in the neurodevelopmental effects of endocrine disrupters were related to thyroid disruption and the thyroid remains relatively poorly assessed in validated guideline studies. Again, examination of serum thyroid hormone and thyroid-stimulating hormone levels following exposure during critical stages of development is only included in the extended one-generation reproduction study (TG 443). Equally, changes in the function of neural or immune tissues may not be marked by distinct changes in their structure but instead by up or down regulation of receptors for hormones and other factors within the tissue, which would be invisible to the histopathological examinations described. Changes in receptor expression could be measured using immunohistochemistry techniques or proteomics in addition to the examinations described. All the animals in the study, including the parental generation, are killed relatively early in their lifecycles. Senescence and old age are not covered, although both the nervous and immune systems change markedly over the course of an individual's life.

With regards to the assessment of obesogens, scientific interest is also very novel and relevant endpoints have yet to find their way in regulatory test methods. Furthermore, changes in body weight have often been used to disregard a potential effect of a chemical; weight loss in animals exposed via their diet could be related to the chemical substance rendering the food simply unpalatable. There are nonetheless other endpoints that may be of interest such as triglycerides and cholesterol levels, but at present these are optional in the 28-day repeated toxicity study (TG 407).

4.7.2 WILDLIFE

4.7.2.1 Invertebrates

A summary of the endocrine related endpoints identified to date in invertebrates is listed in Table 3. Knowledge of invertebrate endocrinology is very incomplete if not non-existent for most phyla. To date, research into invertebrate endocrine systems has centred on molluscs and arthropods, and has been limited in its scope, focusing on reproduction in aquatic molluscs, insects and crustacea and

moulting in insects and crustacea. Compounds inhibiting moulting in insects have been developed as insect control products, and it is possible that these could have effects on both non-target insect populations and crustacea due to close biochemical similarities between the insect and crustacean moulting processes (Annex 1, 7.1.2, 7.1.2.1).

Table 3. Endocrine related endpoints identified in invertebrates (Annex 3, 2.1, Table 28)

Reproduction	Growth and Development	Behaviour
Steroid synthesis/ metabolism	Embryonic, larval and juvenile rates of development	Burrowing behaviour in aquatic molluscs (peppery furrow snail)
Vitellogenin production	Larval and adult survival, body length and weight	
Fertilised egg production and viability	Moult age, frequency and degree of completion	
Offspring viability	Shell/exoskeleton growth and development	
	Sex ratios	
Induction of imposex and intersex conditions	Hormone metabolism	
Time to brood release	Ecdysone levels	
	Retinoic acid pathway function	
Brood sizes		
Mating success	Metabolic disruption (O:N ratio)	

The induction of imposex in female molluscs by organotin antifouling agents is the best documented example of ED in invertebrates in the field. Intersex and oocyte atresia have been observed in marine molluscs living in areas polluted by an oil spill and industrial effluent (Annex 1, 7.1.1.2). Skewed sex ratios and testicular oocytes have been observed in crustacea and vitellogenesis can be induced in mussels near sewage outfalls (Annex 1, 7.1.1.2). The use of juvenile hormone mimics has been suggested to increase crustacean larval mortality, inhibit metamorphosis and reduce their reproductive capacity (Annex 1, 7.1.1.2.). Experimentally, both arthropods and molluscs can be affected by vertebrate sex steroids, adversely affecting larval morphology and adult fertility in arthropods and inhibiting mollusc fertility (Annex 1, 7.1.1.2.). The effects of organotins on molluscan sexual differentiation were initially thought to be the result of enzyme inhibition leading to increased testosterone levels, but more recent work has revealed that it is more likely to be due to the induction of the RXR pathway[1].

Invertebrate endpoints are not covered by the OECD conceptual framework guidance documents on the basis of the lack of diagnostic screening endpoints and little experience in their use. Two assays have recently been validated, the chironomid lifecycle test (TG 233) and the daphnia reproduction

[1] Lima D, Reis-Henriques MA, Silva R, Santos AI, Castro LF, Santos MM. 2010. Tributyltin-induced imposex in marine gastropods involves tissue-specific modulation of the retinoid X receptor. Aquat Toxicol. 2011. 101:1:221-227

test (TG 211). The chironomid lifecycle test lasts 2 generations and measures egg rope and offspring number and viability over both, developmental rate over both, and egg rope fertility in the first generation. The daphnia reproduction test lasts 21 days and measures parental mortality, reproductive output and juvenile survival. Growth measurements can be taken but are considered optional (Annex 3, 7.1, tables 29, 30 and 32). A multigenerational assay utilising mysid shrimps and a full lifecycle test using copepods are also under development. The mysid shrimp test is intended to cover endocrine mediated reproductive endpoints, steroidogenesis and ecdysteroid related endpoints, and the copepod assay to identify estrogenic, androgenic or antiandrogenic properties affecting egg production, and embryo/neonate development (Annex 1, 7.1.1.4., Annex 3, 7.1, tables 29, 30 and 32). These assays do not give a comprehensive view of the effects of EDCs on all sensitive endpoints and are strictly limited to arthropods, mostly crustacea. Thus far, although other assays available in the wider scientific literature using various species of mollusc and insects could be developed to cover more endpoints and effects, this would not close the knowledge gap regarding other invertebrate phyla (Annex 3, 7.1, tables 29, 30 and 32).

In summary, knowledge of invertebrate endocrinology and how it is affected by EDCs is largely confined to arthropods and molluscs. Well documented examples of EDCs affecting populations in the field exist, but so far only endpoints involving reproduction in molluscs and arthropods and moulting in insects have been investigated in detail. Assays using crustacea and molluscs are under development, but other invertebrate phyla have yet to be included.

4.7.2.2 Fish

Since the effects of environmental estrogens on fish were first discovered in the 1970s, several ED sensitive endpoints have been identified that can be used in field and laboratory studies as markers for endocrine disruption. These include sex steroid mediated effects on reproduction and reproductive development (Annex 1, 7.2.1.), and changes in thyroid hormone synthesis and transport, (Annex 1, 7.2.2.2.). In the field, exposure to sewage treatment works effluent has been found to induce intersex and impair reproductive function in male fish (Annex 1, 7.2.1.1.1.). Pulp and paper mill effluents have been found to have similar properties (Annex 1, 7.2.1.1.3), and demasculinisation of primary and secondary reproductive characteristics has been noted in wild populations exposed to runoff from agricultural sites (Annex 1, 7.2.1.1.4.). Fish taken from water courses contaminated with perchlorate have thyroid abnormalities, and some studies have linked PCB exposure with reduced peripheral thyroid hormone deiodinisation (Annex 1, 7.2.2.2.). In the laboratory, ethinylestradiol can be used to enhance egg production in female fish, demasculinise males and alter sex ratios. Females exposed during early development have enhanced sensitivity to estrogens later in life. The natural estrogens estradiol and esterone have similar effects, with estradiol being the most potent (Annex 1, 7.2.1.2.1.). Male and female sexual characteristics can also be manipulated using experimental antiandrogens, androgenic drugs and aromatase inhibitors (Annex 1, 7.2.1.2.2.). There is some evidence that cortisol exposure can cause masculinisation, possibly because it bears a close structural resemblance to 11-keto testosterone, an important androgen in fish (Annex 1, 7.2.2.4.). The effects of chemical exposure on the HPA axis and other non sex steroid and thyroid hormone mediated processes remain insufficiently studied, especially in wild populations. Several assays that can be used to assess the effects of EDCs in fish are contained in the OECD conceptual framework. The short term reproduction assay (TG 229), 21 day fish assay (TG 230) and its variant the androgenised female stickleback assay, and the fish sexual development test

(draft TG 234) measure reproductive endpoints (Annex 1, 7.2.1.4., Annex 3, 7.2 table 35). A multigenerational assay using medaka which will cover both reproductive and some developmental endpoints is also at the draft stage (Annex 3, 7.2 tables 35 and 38). The short term reproduction assay and 21 day fish assays do not assess effects during sensitive developmental times. The fish sexual development test and medaka multigenerational test cover multiple reproductive and development endpoints. Gross organ morphology, time to maturity and thyroid hormone levels are not covered by the OECD conceptual framework, but are included in a test validated by the USEPA, the fish lifecycle toxicity test. Effects on the retinoid system, growth hormone levels or larval metamorphosis were endpoints suggested in the OECD draft Detailed Review Paper on Novel Endpoints and Assays. Changes in behaviour, particularly reproductive behaviour are noted (Annex 3, 2.2. tables 35 and 38).

Effects of EDCs on reproductive endpoints in fish are well documented both in the field and the laboratory. Other endpoints, particularly those not directly involving the HPG and HPT axes have received less attention. Whilst there are assays that available or nearing validation that cover reproductive endpoints, other ED endpoints have been neglected and their inclusion in standardised validated assays should be given due consideration.

4.7.2.3 Amphibians

Intersex and demasculinisation have been reported widely in amphibian populations exposed to industrial and agricultural runoff. Whilst the role EDCs play in this remains controversial, evidence has accrued that the incidence of intersex has increased over time in industrial and agricultural areas, and studies measuring the effects of individual herbicides and herbicide mixtures applied during sensitive stages in the amphibian lifecycle have found associations between their application and elevated thyroid hormone levels, intersex and gonadal dysgenesis in males (Annex 1, 7.3.1.1.). Growth inhibition, elevated thyroid hormone levels and thyroid hyperplasia have been detected in frogs exposed to perchlorate and residues in oil sand extraction sites and metamorphosis has been reported to be delayed or inhibited in frogs exposed to sewage treatment works effluent (Annex 1, 7.3.2.1.). Elevated corticosterone levels have been detected in wild populations living in sites contaminated by coal combustion by-products, artificially increasing corticosterone levels can impair mating behaviour in toads, and a decrease in the time to metamorphosis is a natural response to elevated corticosterone levels increase in tadpoles. Even so, the effects of chemicals on the HPA axis in amphibia have received little attention (Annex 1, 7.3.3.). The experimental administration of natural and synthetic estrogens, androgens and antiestrogenic and anti-androgenic drugs produces alterations in reproductive endpoints similar to those seen in studies of chemical effects (Annex 1, 7.3.1.2.). Thyroid function and metamorphosis have also been demonstrated to be vulnerable to the effects of known EDCs, estrogen, thyroid hormones and their inhibitors (Annex 1, 7.3.2.3., Annex 1, 7.3.2.2.). In the current OECD conceptual framework, the xenopus metamorphosis assay (TG 231) can be used to screen for thyroid disruptors, but this only covers thyroid histopathology and limb development. A larval amphibian growth and development assay is currently in its draft stages and will cover gonad histopathology, vitellogenesis, secondary sexual characteristics and changes in sex ratio, growth, time to metamorphosis and thyroid hormone levels. Further endpoints could be added to the xenopus metamorphosis (TG 231) assay to cover growth and development. Levels of sex hormones have been omitted so far from the conceptual framework and no other endocrine related endpoints have been included in the framework in any form (Annex 7, 2.3, table 41).

To summarise, EDCs have been shown to affect reproductive and thyroid related endpoints in the field and laboratory, but behavioural endpoints have yet to be considered at a regulatory level.

4.7.2.4 Reptiles

Endocrine disruption in reptiles is not studied extensively. Most of the research to date has focused on alligators, with a few studies in turtles. Lizards, which together with snakes make up 96% of reptile species, are represented only by a single study (Annex 1, 7.4). Most of the field evidence so far has been accrued in the aftermath of the contamination of Lake Apopka in Florida by large quantities of persistent organic pollutants. Reproductive and steroidogenic abnormalities were widely reported in the local alligator population, which declined significantly after the spill. Eggs collected from the site and incubated at mixed sex producing temperatures produced instead female biased sex ratios. Studies of alligators at other sites, however, often failed to replicate these results (Annex 1, 7.4.1.1.). Freshwater turtles have been the subject of field studies of populations in sites around the Great Lakes which have suggested that genital sexual dimorphism was reduced in both sexes of free living turtles and their captively hatched offspring from sites polluted with persistent EDCs. Other measures of ED such as serum testosterone and penis morphology were unaffected, and females appeared to be more sensitive than the males (Annex 1, 7.4.1.1.). Sex ratios have been manipulated experimentally in turtles without affecting phallus morphology by administrating estradiol and testosterone *in ovo*. Perhaps unexpectedly, testosterone induced feminisation, presumably due to *in ovo* aromatisation. Treatment with an aromatase inhibitor induced masculinisation. Similar effects on sex ratios, again without effect on phallus morphology or plasma testosterone, have been found in sex steroid or aromatase inhibitor treated alligators. Work done in Japan has determined that vitellogenesis is not a good biomarker of estrogenicity in at least one species of turtle (Annex 1, 7.4.1.2). Bisphenol A, and various PCBs have been demonstrated to feminise alligator sex ratios and reduce egg plasma testosterone levels at male producing temperatures, and atrazine may be able to feminise sex ratios at mixed sex but not male producing temperatures. Similar changes have been seen in estrogenic EDC exposed turtles. Atrazine and endosulfan have been found to disrupt testicular and ovarian morphology in caiman (Annex 1, 7.4.1.3.). Research on the effects of EDCs on non-reproductive endpoints in reptiles is practically non-existent. Elevated T4 but not T3 levels have been found in alligators from a highly contaminated site in Florida. In another site on the same lake thyroid follicle colloid levels were decreased, but a study at a different site showed decreased T4 (Annex 1, 7.4.2.). Behavioural endpoints have received no scrutiny whatsoever in the field or laboratory. The OECD conceptual framework does not cover any aspect of ED effects on reptiles (Annex 3, 7.3, table 41)

Endocrine disruption in reptiles remains a largely unexplored area of research and is not covered by the assays currently validated to guide EDC regulation. The research which exists is limited to very few species and a tiny selection of compounds but demonstrates that population level effects on sex ratios and gonadal histopathology are plausible. Interpretation of these data and identification of ED effects is further hindered by a poor understanding of reptilian endocrinology and the natural variations in hormone levels and body morphology extant in wild populations.

4.7.2.5 Birds

Most work done assessing the effects of EDCs in wild bird populations concerns persistent, mostly legacy compounds such as organochlorine pesticides, PCBs and PBDEs. Skewed sex ratios have been correlated with body burdens of various persistent organic pollutants in seabirds, although it is not clear whether this is an ED effect. Malformations of the reproductive tract, alterations in hormonally induced physiological processes and decreased reproductive performance have been correlated with exposure to PCBs, other chlorinated hydrocarbons and methylmercury compounds in several species (Annex 1, 7.5.1.1.). Thyroid-related impairment of egg development has been correlated with egg polyhalogenated aromatic hydrocarbon content (Annex 1, 7.5.2.1.). Predatory birds at or near the top of their respective food chains in areas polluted with persistent organic pollutants have an increased incidence of thyroid and thyroid hormone abnormalities. Changes in reproductive behaviours have also been noted (Annex 1, 7.5.3.1.). The manipulation of avian reproductive development and behaviour experimentally with sex hormones demonstrates that these endpoints are sensitive to EDCs. The effects of antiestrogens and antiandrogens are less straight forward, since in some studies they have been shown to raise blood levels of the hormones they are supposed to inhibit (Annex 1, 7.5.3.2.). The manipulation of the HPT axis with thyroid hormones, their analogues or antagonists has thus far not been investigated in any detail (Annex 1, 7.5.2.2.). The administration of TCDD, DDT and other estrogenic EDCs *in ovo* has been found to induce malformations of the female reproductive tract and alter adult sex hormone levels, reducing reproductive fitness (Annex 1, 7.5.1.3.). Quite dramatic changes in reproductive behaviour can also be induced, including increased incidences of homosexual pairings (Annex 1, 7.5.3.3.). Effects on thyroid hormone levels and related endpoints have been induced by the administration of PBDEs, TCDD, daidzein and PCBs. Some studies have reported transgenerational effects induced by PBDEs (Annex 1, 7.5.2.3.). The OECD conceptual framework contains 1 validated assay, the avian reproduction test (TG 206) and another, the avian 2 generation test, is at the draft stage. Both assays cover some egg-related reproductive endpoints and gross thyroid pathology, but only the 2 generation assay covers thyroid histology and hormone levels/ related endpoints, sex steroid levels, reproductive tract pathologies, semen quality, adrenal histology and a limited range of courtship behaviours. Corticosteroid levels, sexual preference and behaviours not related to courtship, bone deformities, puberty, offspring survival and yolk retention are not covered by either assay (Annex 3, 7.4, tables 44, 45 and 47).

In conclusion, field studies of the effects of EDCs on wild birds have focused on mostly persistent chemicals. Abnormalities of the reproductive tract, thyroid function and hormonally sensitive behavioural endpoints have been reported in the wild and can be induced in the laboratory with model EDCs and hormones. The validation and adoption of the avian 2 generation test will improve the range of endpoints covered, but non EAT, and many behavioural and developmental endpoints still need to be included.

4.7.2.6 Mammals

Marine mammals have been found to be particularly at risk from the effects of persistent EDCs. Like humans, they have long life spans and occupy high trophic levels. In addition they have large reserves of body fat which act as sinks for lipophilic compounds which are then released in quantity during pregnancy, lactation and times of famine. They are however difficult to study in the wild and often impossible to study in captivity due to their aquatic habitat, long and often poorly documented

migrations and in some cases large body sizes and rarity. High body burdens of PCBs and other POPs have frequently been recorded, sometimes at levels known to cause reproductive abnormalities in other marine or terrestrial species. Pinnipeds have received more attention than cetaceans and due to their relatively small size and abundance have been utilised in some feeding studies. In wild populations reproductive abnormalities and failure have been reported in areas polluted with POPs and in some cases correlated with contaminant body burdens. Seals experimentally fed fish from contaminated areas had increased rates of reproductive failure. Thyroid abnormalities, bone lesions and lowered serum thyroid hormone and retinol levels have been found to correlate with contamination in wild populations. High rates of reproductive failure, increased CYP1A1 expression, and ovarian cysts indicating impeded ovulation have been observed in cetacean populations feeding in contaminated areas, and in some of these areas population recovery from hunting has been slower than expected or has failed altogether. Mice fed diets including blubber from heavily contaminated animals developed reproductive abnormalities. Lowered blood serum thyroid hormone levels have also been correlated with POP body burdens in cetaceans, and body length has been found to relate to burdens of certain PCB congeners in a sex specific manner in one study of bowhead whales. Thus far, studies on the effects on non-EAT mediated endpoints have not been undertaken in marine cetaceans. Adrenal lesions related to a Cushing's disease-like phenomenon started appearing in wild seal populations exposed to POPs after WWII and has been found to correlate with DDT and PCB methyl sulphone exposure. Other non-EAT effects have not been investigated, and in both pinnipeds and cetaceans the potential effects on non-persistent or pseudopersistent contaminants remain unknown. Due to the impracticalities of keeping these animals in a laboratory environment, studies utilising model EDCs or hormones are lacking. None of the assays in the OECD framework cover endpoints in marine mammals, those intended to cover human relevant endpoints are assumed to protect marine mammals and other wild mammal species as well. Rodent studies, however, do not offer any insight into the effects of EDCs during old age in a long lived species, the effects of the metabolites which result from biotransformation in blubber.

Endocrine disruption in marine mammals has not been studied in great detail, but there are strong indications that endocrine related endpoints have been affected by POPs in wild populations, possibly to the detriment of their survival. Non EAT endpoints and the effects of non-persistent or pseudopersistent compounds remain unstudied in most cases.

4.8 CHEMICALS OF CONCERN AND EXPOSURE

The last 10 years have seen increases in the number of chemicals considered of concern in the context of endocrine disruption, and in the number of endpoints they have been connected with. Their coverage, however, has been skewed strongly towards reproductive endpoints governed by the actions of the sex steroids, and a large portion of the literature is devoted to legacy compounds such as DDT and its metabolites and PCBs, which are no longer in widespread use.

Table 4 overleaf gives a summary of which chemical groups have been investigated and the endpoints they have been investigated in connection with.

Table 4. Table relating groups of chemicals of concern to human health/wildlife endpoints

Chemicals of concern	Investigated in connection with...																		
	Human Health Endpoints													**Wildlife Endpoints**					
	Male reproductive health	Female precocious puberty	Female fecundity	Polycystic ovary syndrome	Female fertility	Endometriosis	Uterine fibroids	Breast cancer	Prostate cancer	Testis cancer	Thyroid cancer	Developmental neurotoxicity	Metabolic syndrome	Invertebrates	Fish	Amphibians	Reptiles	Birds	Mammals
PCBs, PCDDs, PCDFs*	●	●	●		●	●	●	●	●	●	●	●	●		●	●	●	●	●
Polybrominated biphenylethers (PBDEs)	●	●					●			●		●	●			●		●	●
Perfluorinated compounds (PFCs)			●									●	●		●	●		●	
DDT/DDE	●	●	●		●	●	●	●	●	●		●	●		●	●	●	●	●
Other organochlorine pesticides	●		●		●	●		●		●	●	●	●		●	●	●	●	●
Organo-phosphate pesticides					●				●			●			●	●			
Carbamate pesticides					●				●		●			●	●				
Azole pesticides	●										●								
Pyrethroid pesticides									●										
Triazine herbicides															●	●	●	●	
Other Pesticides	●		●		●				●		●	●		●	●	●		●	
Heavy metals	●	●	●		●			●	●			●						●	●
Alkylphenols, bisphenol A, parabens,		●		●	●	●		●			●	●	●	●	●	●	●		
Phthalates	●	●			●	●	●			●		●	●		●	●			
Pharmaceutical estrogens	●				●	●	●	●	●	●	●	●	●	●	●			●	
Phytoestrogens		●	●			●	●	●	●		●	●						●	
Organotins													●	●	●				

*Polychlorinated biphenyls (PCBs), dioxins (PCDDs), furans (PCDFs)

4.8.1 POLYCHLORINATED BIPHENYLS (PCBS)

Evidence of exposure-disease associations

Decreased sperm quality has been noted in humans exposed to PCBs both *in utero* and in adulthood. Tentative links have been made between high PCB body burdens and irreversible reproductive endpoints such as precocious female puberty, pregnancy loss and low birth weights (Annex 1, 4.4.3., 4.5.3.3). There is also some evidence that they may be associated with skewed offspring sex ratios, but the results of studies of this endpoint are conflicting (Annex 1, 4.5.3.3). A convincing link has been made between PCB exposure and an increased risk of breast cancer in women with certain CYP polymorphisms, and there is some evidence that the incidences of fibroids, thyroid cancer and prostate cancer are also increased (Annex 1, 4.7.3., 5.4.2.4., 4.5.3.3) respectively. Neurodevelopmental endpoints are also known targets of PCB exposure, which has been linked to serious and irreversible effects on cognition, motor and sensory function (Annex 1, 6.1.3). Estrogen-related developmental perturbations, thyroid irregularities and the suppression of thyroid hormone synthesis have also been found in a variety of exposed vertebrate wildlife species (Annex 1, 7.2.1.1.2, 7.3.1.1, 7.4.1.1, 7.5.1.1, and 7.6.1.1).

Evidence of endocrine disruption mechanisms

Briefly, PCBs are known to target various endocrine and neuroendocrine mediated endpoints via effects on a broad profile of receptors and enzymes. The precise targets vary from compound to compound and depend on the coplanarity of congeners. For co-planar PCBs and their hydroxylated metabolites they include the thyroid receptors and HPT axis[1], for non-coplanar PCBs (and their hydroxylated metabolites) the estrogen and androgen receptors and HPG axis, the AhR and *CYP1A1* induced steroid metabolism[2].

4.8.2 POLYCHLORINATED DIBENZODIOXINS (PCDDS) AND POLYCHLORINATED DIBENZOFURANS (PCDFS)

PCDDs and PCDFs share some of the same mechanisms of action as co-planar PCBs, binding to the AhR, perturbing thyroid function, but they do not activate the estrogen receptor. Exposure to TCDD during infancy led to irreversible reductions in sperm motility and sperm concentration. Strikingly, the opposite effect was observed among men who were exposed during puberty (Annex 1, 4.1.3.3.2). They may also irreversibly affect female reproductive endpoints and neurodevelopment in both genders. Delays in female puberty have been associated with pre- and perinatal exposure, (Annex 1, 4.2.3) and PCDD/PCDF can adversely affect neurodevelopment in a similar manner to PCBs (Annex 1, 6.1.3.1.1.2). Reduced age at menopause has been associated with exposure in adult women (Annex 1, 4.3.3.2). There is an association with high PCDD/PCDF exposures and breast cancer, and a suggestive, almost significant association with thyroid cancer (Annex 1, 5.1.3.2). (Annex 1, 7.5.1.3.).

[1] Patrick L. 2009. Thyroid disruption: mechanism and clinical implications in human health. Altern Med Rev 14:326-346.

[2] Bradshaw TD, Bell DR. 2009. Relevance of the aryl hydrocarbon receptor (AhR) for clinical toxicology. Clin Toxicol (Phila) 47:632-642

4.8.3 POLYBROMINATED BIPHENYLETHERS (PBDES)

Evidence of exposure-disease associations

Exposure *in utero* or during early development can have profound and irreversible effects on neurodevelopment. In humans it has been linked to IQ deficits and there is evidence suggesting that exposure may be a risk factor for autism, although this is still under investigation[1]. There is also evidence from human epidemiological studies that PBDEs might cause reproductive disturbances. There is some evidence that human exposure *in utero* may cause irreversible histological changes in the testis which increase the risk of testicular germ cell cancer (Annex 1, 5.3.3.2.), and associations have been found between body burdens of some compounds and uterine fibroids in women (Annex 1, 4.7.3.). Metabolic endpoints may also be affected, with recent epidemiological studies showing an association between body burdens of PBDEs and an increased risk of diabetes and metabolic syndrome[2]. The effects of these compounds in wildlife are not well studied. Increased thyroid hormone levels have also been correlated with PBDE body burdens in wild seal pups (Annex 1, 7.6.3.1.).

Evidence of endocrine disruption mechanisms

Mechanistic studies suggest that as well as binding to the thyroid receptors and the thyroid hormone and retinol transport protein transthyretin, PBDEs also act as EDC by inducing the enzymes involved in glucuronidation, downregulation of the transport protein transthyretin, or downregulation of thyroid hormone transport. Experimental administration of low doses can cause thyroid abnormalities, delay metamorphosis in amphibians (Annex 1, 7.3.2.3.), impair avian thyroid hormone synthesis, and reduce egg quality and thyroid hormone content (Annex 1, 7.5.1.3. 7.5.2.1.). It can also cause irreversible behavioural changes, decreasing courtship and other reproductive behaviour. Their action on transthyretin also makes them potent disruptors of retinoic acid signalling. An extensive body of work carried out *in vivo* and *in vitro* supports the epidemiological evidence that exposure during critical periods of neurodevelopment has an adverse effect on neurogenesis and brain development (Annex 1, 6.1.3.1.2.). Effects on neurodevelopment are thought to be at least partly caused by these compounds thyroid disrupting properties, but their ability to inhibit dopamine uptake and produce free radicals may play a role too. They also act as AR antagonists and ER agonists, and slow estrogen clearance by inhibiting the activity of sulfotransferases. Ovarian abnormalities can be induced in chicks dosed experimentally with TCDD, and in adulthood they have lower progesterone levels and lay smaller eggs. Evidence from animal models suggest that these compounds may have a delaying effect on puberty (Annex 1, 4.2.3.) and may be toxic to the developing testes (Annex 1, 4.1.3.1.4.), although there are few data relating to their effects on these endpoints in humans and those that exist are far from conclusive.

[1] Messer A. 2010. Mini-review: Polybrominated diphenyl ether (PBDE) flame retardants as potential autism risk factors. Physiology & Behavior 100:245-249

[2] Bruggeman V, Onagbesan O, Dumez L, De Ketelaere B, Decuypere E. 2005. Effects of early prenatal exposure to 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) on postnatal reproduction in the laying hen (*Gallus gallus*). Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology 141:349-355.

4.8.4 PERFLUORINATED COMPOUNDS (PFCS)

Evidence of exposure-disease associations

In humans, links have been made between PFC blood serum levels and ADHD in children (Annex 1, 6.1.3.1.6.). In a human study of associations between perfluorooctanoic acid (PFOA) and thyroid hormones a positive association with T3 and a negative association with free T4 was found suggesting that these chemicals interfere with thyroid hormone conversion (Annex 1, 6.1.3.6). Although a causal link has yet to be established, mechanistic and epidemiological studies show a consistent positive association between PFC body burdens and increased cholesterol levels, indicating that they could have an effect on the aetiology of metabolic syndrome and other metabolic disorders (Annex 1, 6.2.3.2.2). The strength of this association however remains controversial. Reproductive endpoints may also be affected. Increases in time to pregnancy following a positive dose-response pattern have been detected in a large epidemiological study (Annex 1, 4.3.3.2.), and high maternal body burdens have been associated with low birth weights. The effects of PFCs on wildlife have yet to be ascertained.

Evidence of endocrine disruption mechanisms

The main classes of concern are the perfluorocarboxylic acids, of which PFOA has received the most attention, perfluorinated phosphinates and alkylphosphonates, and perfluorocyclohexane sulfonates. All these compounds are known to inhibit thyroid function, reducing blood serum thyroid hormone levels *in vivo* in both humans and experimental animals, including both mammals and fish. Experiments conducted *in vitro* and *in vivo* support their status as developmental neurotoxicants, demonstrating that they are capable of causing irreversible neurological changes *in utero*. In rodents, animals exposed *in utero* show delayed perinatal lung development and a slower rate of growth and development after birth. In female rats and their pups decreases in T3 levels were found following short or long term exposure (Annex 1, 6.1.3.6). Some studies of its effects on metabolism suggest it might affect insulin levels via the PPARα and RXR, but this is controversial (Annex 1, 6.2.3.2.2).

4.8.5 PESTICIDES

The organochlorine pesticides, including DDT and its metabolites and HCH were amongst the first compounds to be investigated for their endocrine disrupting properties, and other classes are known or suspected to have endocrine disrupting properties too, but here three classes of compounds which have been the subject of increasing attention over the last 10 years are discussed.

4.8.5.1 Dicarboxamides

Evidence of exposure-disease associations

No direct evidence of exposure-disease associations exists yet for either humans or wildlife.

Evidence of endocrine disruption mechanisms

Results of animal studies indicate that vinclozolin could contribute to male reproductive disorders. *In vivo*, prenatal vinclozolin exposure causes irreversible adverse effects on male sexual differentiation

in the rat. When pups are exposed *in utero* during sexual differentiation, androgen dependent processes are inhibited causing AGD decreases and genital deformities. In adulthood the exposed males have testicular abnormalities and poorer spermatogenesis, prostate inflammation and disease, and breast tumours, along with kidney disease and immune abnormalities. The principal dicarboxamide pesticides to receive attention are the fungicides vinclozolin, iprodione and procymidone. The antiandrogenic properties of vinclozolin in particular have been a cause for concern, since there are indications that its prenatal effects on the male reproductive tract may persist for two or more generations (Annex 1, 4.1.4.). Most of the effects ascribed to vinclozolin can be attributed to its properties as an androgen receptor antagonist (Annex 1, 4.1.4.) but in addition to these properties, it causes genome-wide gene methylation changes which may allow these effects to be passed on transgenerationally, although the exact mechanism of action at the genetic level has yet to be determined (Annex 1, 3.4.3.1.). Vinclozolin has also been shown to alter progesterone receptor expression *in vivo*, having a virilising effect on female mice. Iprodione and procymidone are also AR antagonists, causing male reproductive abnormalities similar to vinclozolin, although it is not yet known whether these are passed to subsequent generations (Annex 1, 4.1.4.).

In wildlife species, vinclozolin can impair reproductive function in male prosobrach molluscs at environmentally relevant concentrations (Annex 1, 7.1.1.4.). It can also inhibit maturation and reduce egg production in fish, females being more sensitive than males possibly because they have far lower endogenous androgen titres and because androgens play an essential role in folliculogenesis (Annex 1, 7.2.1.3.4.). Masculine behaviour is also inhibited in birds exposed *in ovo* and adult males show reduced blood testosterone levels and have smaller testes (Annex 1, 7.5.3.3), but experiments exposing reptiles *in ovo* have failed to find any difference in plasma sex steroid levels or sex ratios (Annex 1, 7.4.1.2.).

4.8.5.2 Azole fungicides

Evidence of exposure-disease associations

Azole fungicides can be broadly broken into two categories, triazoles and imidazoles. Prochloraz is the most thoroughly researched imidazole with ED properties, whilst the triazole category includes bitertanol, cyproconazole, febuconazole, epoxiconazole, hexaconazole, metconazole and myclobutanil. Neither they nor the dicarboxymides are persistent, so measuring or estimating exposure at critical times is difficult, especially in humans. The activities of these chemicals in humans have received little epidemiological attention. Studies of agricultural workers have detected statistically significant increases in reproductive abnormalities in the children of mothers exposed to pesticides during pregnancy. For example, one very recent study from Denmark found an increase in the incidence of cryptorchidism in the sons of a cohort of mothers employed in horticulture and exposed to pesticides relative to cohorts of unexposed mothers employed both in horticulture and non-agricultural occupations[1] Unfortunately, these studies do not identify the individual compounds or compound classes that the women were exposed to, so causal links between individual pesticides or pesticide classes cannot be inferred from these data. Little is known about how dicarboximides may affect wildlife populations.

[1] Gabel P, Jensen MA, Andersen HR, Baelum J, Thulstrup AM, Bonde JP, Toft G. 2011 The risk of cryptorchidism among sons of women working in horticulture in Denmark: a cohort study. Environmental Health 10:100-

Evidence of endocrine disruption mechanisms

Exposure to prochloraz during organogenesis interferes irreversibly with the formation of the male reproductive tract in a similar manner to other antiandrogens, and there is evidence that other azoles act similarly (Annex 1, 4.1.4.). In addition to being an AR antagonist, prochloraz and the triazoles also interfere with steroid synthesis. In the case of prochloraz, it involves the inhibition the conversion of progesterone to testosterone, the stimulation of aldosterone synthesis at low doses and inhibition of it at high ones. It is also an aromatase inhibitor, allowing it to reduce estradiol levels as well. Other azoles show similar properties, but their individual abilities to inhibit the conversion of progesterone and the action of aromatase vary (4.1.2.). Non-steroidal activity has also been confirmed. Azoles can also enhance the hepatic metabolism and excretion of the thyroid hormones (Annex 1, 4.1.4). Quite a few of the azole fungicides have been used in mixture experiments, and appear to have additive effects in line with dose addition predictions (Annex 1, 3.3.1.2.). Although they are known to affect steroidogenesis in amphibians in laboratory studies (Annex 1, 7.3.1.3.), it is unknown whether they affect them in the field. Fish may be at risk from these compounds, since laboratory studies report reduced fecundity and gonadal abnormalities in individuals exposed during adulthood, and in fish exposed during development increases in the male:female ratio, an increased incidence of intersex and increased vitellogenesis in males were recorded [1,2]. Data for other wildlife groups are lacking.

4.8.5.3 Triazines

Evidence of exposure-disease associations

Atrazine and simazine are the two most widely used herbicidal triazines, and of these atrazine has received the greatest attention for its ED properties, principally because feminised secondary sexual characteristics, intersex and gonadal dysgenesis have been observed in wild frogs collected from contaminated sites (Annex 1, 7.1.1.4.).

Evidence of endocrine disruption mechanisms

Exposure is associated with lower testosterone levels in amphibia, and in female rats, peripubertal exposure to atrazine reduces luteinising hormone (LH) and prolactin levels, delaying pubertal onset. Its exact mechanism of action, however, is not known. It is not known to interact with any hormone receptor directly, but interferes with steroidogenic enzymes (Annex 1, 7.3.1.3.). Surprisingly, it does not appear to affect aromatase. Similar sexual abnormalities to those seen in male frogs in the wild have been induced experimentally, with some studies using very low doses of atrazine (Annex 1, 7.2.1.3.4.). The effects observed, however, differ between species and strains. Crustacea may have some reproductive vulnerability to triazines (Annex 1, 7.3.1.3.). When exposed to atrazine under experimental conditions, egg bearing water fleas produced a disproportionate number of male offspring. Some studies using fish have found elevated vitellogenin production in exposed males,

[1] Ankley GA, Jensen KM, Durhan EJ, Makynen EA, Butterworth BC, Kahl MD, Villeneuve DL, Linnum A, Gray LE, Cardon M, Wilson VS. 2005 Effects of Two Fungicides with Multiple Modes of Action on Reproductive Endocrine Function in the Fathead Minnow (*Pimephales promelas*). Toxicological Sciences 86(2):300-308

[2] Kinnberg K, Holbech H, Petersen GI, Bjerregaard P. 2006 Effects of the fungicide prochloraz on the sexual development of zebrafish (*Danio rerio*). Comparative Biochemistry and Physiology Part C: Toxicology and Pharmacology 145(2): 165-170

and reduced egg production and numbers of spawning events in females (Annex 1, 7.2.1.3.4.). In these studies, testicular oocytes, oocyte atresia and other histological changes were observed in the gonads of treated fish, although not in a dose dependent manner. Little is known about how these compounds affect the endocrine systems of wild reptiles, birds and mammals. Preliminary studies of effects on the sex ratios of alligator eggs found no effects at male or female producing temperatures, but the temperatures producing mixed sex clutches, which are the most sensitive, were not tested (Annex 1, 7.4.1.3.). Male courtship behaviour and the formation of the cloacal gland can be suppressed in birds by the administration of atrazine *in ovo* (Annex 1, 7.5.3.3.). Effects in mammals remain largely undetermined.

4.8.6 HEAVY METALS

Evidence of exposure-disease associations

Due to the persistent and bioaccumulative nature of these compounds they are still a cause for concern. The irreversible neurodevelopmental effects of methylmercury are currently a topic of great interest since epidemiological studies have linked exposure to delays in important developmental milestones and deficits in cognitive, motor, auditory and visual function. (Annex 1, 6.1.3.). Lead has long been known to perturb these endpoints and its use in many products has been restricted as a result. The reproductive effects of cadmium are also receiving scrutiny. Some epidemiological studies of occupational exposure make a weak link between cadmium exposure and breast cancer (Annex 1, 5.1.3.2.). There is also some evidence that it may contribute to menstrual abnormalities and increased time to pregnancy (Annex 1, 4.3.3.2), but no association has been found with endometriosis (Annex 1, 4.6.3.) and studies examining potential links between exposure and the incidence of fibroids give conflicting results (Annex 1, 4.7.3.). Weak associations have been made between exposure and prostate cancer, although these associations tend to be stronger for more aggressive forms of the disease. Its ED effects in wildlife are understudied. Although the bioaccumulation of cadmium has been widely reported in populations of vertebrate and invertebrate wildlife[1], and effects on individual bodily processes have been recorded in a variety of species both experimentally and in the wild, data on population level effects are lacking.

Evidence of endocrine disruption mechanisms

Methylmercury has multiple modes of action relating to the endocrine system. It perturbs thyroid hormone homeostasis by increasing TSH levels, decreasing T3 and increasing T4 levels, and can in addition act as a toxicant to the endocrine organs directly, interfere with steroidogenesis, interact with the sex hormones directly and change hormone concentrations[1]. Like methyl mercury, lead, long known to be a neurotoxicant, exerts an endocrine mediated action via the enhanced pituitary release of TSH. Investigations using the uterotrophic assay show a dose-response relationship between cadmium exposure and uterine weight. There is evidence supporting the role of cadmium as a prostate carcinogen in rodents (Annex 1, 5.2.3.3.). Detailed investigations of the effects of cadmium on the ERα in animals show it to behave very similarly to estrogen, and most of the studies carried out *in vivo* have proven to be reproducible. *In vitro*, however, there is more confusion. Proliferative effects on estrogen sensitive cell cultures have been reported by some studies but not

[1] Burger J. 2008. Assessment and management of risk to wildlife from cadmium. Science of the Total Environment 389(1):37-45

by others, similarly, studies of cadmium induced ERα and transcription have produced contradictory results, and studies showing that cadmium exposure induced ERα mediated phosphorylations of the Erk1,2 kinases in HEK293, HeLa, MCF7 and HepG2 cells could not be replicated in some laboratories. Results gleaned from studies of cadmium and estradiol mostly demonstrating that cadmium sensitises the ERα to estradiol action, although some have found the opposite.

4.8.7 BISPHENOL A

Bisphenol A has been researched very thoroughly over the last few years. Its effects are multifaceted, mediated by its ability to bind the ER and PR, and its properties as a thyroid hormone antagonist. Exposure during organogenesis has been demonstrated to have irreversible adverse effects on reproductive development, namely the hyperplasia of the prostate gland and increases in its sensitivity to estrogen later in life (Annex 1, 3.2.4.1.1.), and changes in the histoarchitecture of mammary tissue (Annex 1, 3.2.4.1.2.). Emerging areas of study are the potential cancer risks and adverse neurodevelopmental outcomes which may be associated with bisphenol A exposure (Annex 1, 3.2.4.1.3).

4.8.8 PHTHALATES

Evidence of exposure-disease associations

Two key studies conducted in the USA among young boys provide good evidence of associations of irreversible effects in the form of altered hallmarks of sexual differentiation with elevated phthalate exposures during fetal life. A summary score of urinary phthalate metabolite levels showed associations with shorter anogenital index (AGI). The relationships of shorter AGI with other health outcomes, including testicular descent and genital morphology were also investigated. The likelihood of incompletely descended testes was strongly related to shorter AGI, as was the proportion of boys with a scrotum categorized as small and with a small penis size. Neurodevelopment and metabolic endpoints are emerging areas of concern in relation to phthalate exposure, since studies of prenatal exposure have found associations with phthalate exposure and ADHD-link symptoms and lowered IQs (Annex 1, 6.1.3.)., and exposure has been implied as a risk factor for obesity, insulin resistance and diabetes by others (Annex 1, 6.2.3.2.6.). Effects on wildlife remain largely uninvestigated.

Evidence of endocrine disruption mechanisms

When given to pregnant rats in controlled experimental exposures, phthalates produce a series of irreversible effects in the male offspring, termed the “phthalate syndrome”. This syndrome is thought to have similarities with the testicular dysgenesis syndrome (TDS) in humans (see Annex 1, 4.2.). In rats, it is characterised by a series of reproductive tract abnormalities including underdeveloped reproductive organs, malformations of the external genitalia (similar to hypospadias), difficulties with testes descent (cryptorchidism), changes in anogenital distance and retained nipples. These effects can be traced to disruption of androgen action in fetal life, a clear endocrine mechanism.

The effects on male development are attributed to the ability of phthalates to directly interfere with testosterone synthesis by blocking the uptake of testosterone precursors into Leydig cells. The

capability to suppress androgen synthesis appears to be largely confined to phthalates with ortho configuration and ester side chains between four and six carbon atoms. Shorter chain phthalates, such as diethylphthalates, are inactive in rats when administered orally (Annex 1, 4.1.2.). There are some indications that diethylphthalates may be estrogenic in fish and frogs (Annex 1, 7.2.1.3.2., 7.3.1.3.).

A mechanism analogous to that in male rat offspring is operating in female rats. Mono-ethyl-hexyl phthalate (MEHP), the active metabolite of DEHP, can decrease estradiol production in rat granulosa cells *in vitro*, essentially by interfering with the production and metabolism of the hormone. DEHP can also decrease estradiol production by reducing the levels of aromatase, the enzyme that converts testosterone to estradiol. Some phthalates, including benzyl butyl phthalate (BBP) and DEHP[1] are also ER agonists. Thyroid hormone function can become perturbed, both at the level of the thyroid gland, which has been found to undergo histopathological changes and release less T4 in rodent studies, and in the thyroid hormone-responsive tissues themselves, where phthalates can bind to and suppress the function of the T3 receptors. Some also show an affinity for the GR, and can promote adipogenesis *in vitro*. Some phthalates share structural similarities with known COX-1 and COX-2 inhibitors such as aspirin, raising the possibility that they may also interfere with prostaglandin synthesis (Annex 1, 3.5.2.2.). A mixture of phthalates tested for antiandrogenic properties *in vivo* acted together in line with the dose addition model (Annex 1, 3.3.1.2.).

Experimental approaches, data on potency

Typically, suppression of testosterone synthesis, retained nipples and changes in AGD are the most sensitive effects induced by certain phthalates in the rat, after gestational exposures. These appeared in the absence of maternal toxicity, or any other signs of systemic effects. At higher dosages, malformations of the genitalia and other effect on sex organs were seen.

DEHP appears to be more potent than DBP and BBP, with DINP being the least potent of the phthalates. Based on landmarks of male sexual differentiation, the NOAEL for DEHP is 3-5 mg/kg d, for DBP and BBP around 50 mg/kg d and for DINP 300 mg/kg d.

4.8.9 PARABENS

Evidence of exposure-disease associations

Parabens are emerging as compounds of concern, having been widely reported as contaminants in human tissues and bodily fluids. Most work on them to date has been conducted *in vitro* or using animals. It should be noted, especially for the *in vitro* studies, that parabens are metabolised very quickly by the body, so their as yet untested metabolites are likely to play a role in any endocrine disruption consequent to exposure. Epidemiological evidence in humans is very limited.

An association emerged recently between blood serum paraben levels and mammographic breast density in postmenopausal women, but there is no evidence that exposure may increase breast cancer risk.

Evidence of endocrine disruption mechanisms

In vitro, parabens have been shown to have estrogenic and antiandrogenic properties in receptor binding studies[2]. Their estrogenic properties increase with increasing linear alkyl chain length or branching, and their antiandrogenic properties may follow the same pattern. In adult rats paraben administration has been shown to reduce testosterone levels in a dose dependent manner and reduce sperm production. Increased prostate weights have been reported. Some studies of exposure pre- and perinatally have reported irreversible adverse effects on male development including reduced sperm counts and reduced testis and prostate weights, but the results of the few extant studies are conflicting. Mixtures of parabens have been shown to have additive estrogenic effects. They can also inhibit sulfotransferases, inhibiting the excretion of endogenous estrogens. They have also been found to be capable of TR antagonism *in vitro*, raising concerns about their possible irreversible effects on neurodevelopment, but so far this avenue of inquiry remains relatively unexplored (Annex 1, 6.1.2.1). Like phthalates, parabens may also act as COX inhibitors. Some have been shown to inhibit prostaglandin synthesis *in vitro*, with COX inhibition as a probable mode of action[3]. Studies examining the effects of parabens in mixtures *in vivo* or their effects on wildlife species are largely absent. Overall our knowledge of the ED properties of parabens is inadequate, and much more work is needed to determine whether they pose any risks to human or environmental health.

4.8.10 OTHER CHEMICALS

Multiple new chemicals and groups of chemicals have come to the fore over the last decade as being of potential concern. In addition to parabens, UV filters and artificial musks are present in many cosmetic and personal care products. As of yet direct evidence of either group of compounds having ED effects on humans is lacking, but some evidence exists from studies conducted in animals and *in vitro* that some of them may have endocrine disrupting properties. UV filters are a diverse group of compounds. When administered orally in high doses to rats during pregnancy, some of these compounds have been found to induce irreversible estrogenic and antithyroid effects in the offspring. Estrogenic properties have also been observed in gene reporter assays, and most of the compounds demonstrated to be estrogenic also have AR, TR and/or PR antagonistic properties[1] (Annex 1, 3.1.3.3), although the profiles of activity observed vary widely between compounds. Synthetic musks have been found to accumulate in the tissues of humans and wildlife. Some show activity *in vitro*, acting as very weak estrogenic, antiandrogenic and antiprogesterone compounds. Very few studies have been carried out looking at their behaviour *in vivo*, however. Siloxanes are silicone based compounds frequently used in cosmetics to soften, smooth and moisten skin and hair. Some members of a subcategory of these, the cyclosiloxanes are persistent and have shown the potential to bioaccumulate. They have been found in human tissues as well as in wildlife, but to date there are no data for effects in humans. At high doses some of these compounds have estrogenic effects in rodent assays, and irreversibly impair male reproductive function when administered to pregnant females in non-human primate and rabbit assays. Changes in the reproductive organs of male dogs have also been reported. *In vitro* they have an affinity for the estrogen receptor[1]. Triclocarban and triclosan, which are used widely in antimicrobial soaps, provoke weak aryl hydrocarbon, estrogen and androgen receptor mediated responses in *in vitro* assays and they may also perturb thyroid hormone function (Annex 1, 6.1.2.1). The possibility has also been raised that

triclocarban may interact with testosterone, and *in vivo* there are indications that it can induce hyperplasia of the accessory sex organs of young male rats, but human data are lacking.

Glycol ethers, used as solvents in dry cleaning, and their metabolites have been associated with reproductive toxicity in studies of occupational exposure, namely spontaneous abortion and subfertility in women and decreased sperm counts in men. These findings are supported by animal studies, and an enhancement of estrogen signalling has been noted *in vitro*. All of the compounds above, like most EDCs investigated to date primarily appear to act via disruption of estrogen, androgen or thyroid mediated processes. There are others, however, that act via different pathways which are coming to the attention of investigators as being of concern for human health. Organotin compounds, for example, long known to cause imposex and intersex conditions in molluscs, have been found to promote adipogenesis in mammalian cells via the Peroxisome Proliferator-Activated Receptor γ/Retinoid X Receptor Pathway (Annex 1, 6.2.3.2.4). The identification and quantification of the effects of compounds which affect non-CMR endpoints such as metabolism, bone health, the neuroimmune system and neurodevelopment will be important areas of research in coming years.

5 EUROPEAN REGULATORY FRAMEWORK

5.1 CLASSIFICATION AND LABELLING

The Globally Harmonised System for the Classification and Labelling of Chemicals (GHS) "Purple Book" is an international system for the classification of chemicals drafted under the auspices of the United Nations and was published in 2007. Regulation No 1272/2008 on classification, labelling and packaging of substances and mixtures (CLP thereafter) transposes this classification scheme into European legislation. The CLP amends and repeals the "Dangerous Substances" Directive (67/548/EEC) and the Dangerous Preparations Directive (99/45/EC), and amends the REACH Regulation. There is no specific class for endocrine disrupters within this classification scheme. In order to assess the regulatory basis for 'equivalent concern' for endocrine disrupters as well as the proposals regarding regulatory decisions on EDCs, the current practice for the classification of carcinogens, reproductive toxicants, harmful and toxic substances as well as for environmental toxicants is briefly presented in this section.

It should be noted that in accordance with article 13 of the CLP Regulation, the intrinsic properties of a substance may warrant classification in "one or more" hazard classes.

5.1.1 CARCINOGENS

5.1.1.1 Definition and categorisation

A carcinogen is defined as:

"a substance or a mixture of substances which induce cancer or increase its incidence. Substances which have induced benign and malignant tumours in well performed experimental studies on animals are considered also to be presumed or suspected human carcinogens unless there is strong evidence that the mechanism of tumour formation is not relevant for humans."

Carcinogens are classified under two main categories; category 1, known or presumed carcinogens or category 2 suspected carcinogens. Category 1 is subdivided into categories 1A (known carcinogen) and 1B (presumed carcinogen) primarily on the basis of whether evidence of carcinogenicity is derived from human or animal studies, respectively. The classification scheme and type or strength of evidence associated with each category is reproduced in Table 5.

Among the important factors to consider when assessing the overall level of concern, the following may also be relevant for endocrine disrupters:

"(f) whether responses are in a single species or several species;
(g) structural similarity to a substance(s) for which there is good evidence (of carcinogenicity);
(h) routes of exposure;
(i) comparison of absorption, distribution, metabolism and excretion between test animals and humans;
(j) the possibility of a confounding effect of excessive toxicity at test doses;
(k) mode of action and its relevance for humans (...)"

Table 5. Categorisation of Carcinogens

Categorisation	Definition	Criteria	Test methods
Category 1	Known or presumed human carcinogens A substance is classified in Category 1 for carcinogenicity on the basis of epidemiological and/or animal data.	Category 1A Known to have carcinogenic potential for humans, classification is largely based on human evidence	Human studies that establish a causal relationship between human exposure to a substance and the development of cancer (known human carcinogen)
		Category 1B Presumed to have carcinogenic potential for humans, classification is largely based on animal evidence.	Animal experiments for which there is sufficient evidence to demonstrate animal carcinogenicity (presumed human carcinogen). In addition, on a case-by-case basis, scientific judgment may warrant a decision of presumed human carcinogenicity derived from studies showing limited evidence of carcinogenicity in humans together with limited evidence of carcinogenicity in experimental animals.
Category 2	Suspected human carcinogens	The placing of a substance in Category 2 is done on the basis of evidence obtained from human and/or animal studies, but which is not sufficiently convincing to place the substance in Category 1A or 1B, based on strength of evidence together with additional considerations. Such evidence may be derived either from limited evidence of carcinogenicity in human studies or from limited evidence of carcinogenicity in animal studies.	

5.1.1.2 Sufficient and limited evidence

In relation to the strength of the evidence, the CLP offers definitions for both sufficient and limited evidence that are consistent with IARC definitions for the same terms:

*"**sufficient evidence of carcinogenicity**: a causal relationship has been established between the agent and an increased incidence of malignant neoplasms or of an appropriate combination of benign and malignant neoplasms in (a) two or more species of animals or (b) two or more independent studies in one species carried out at different times or in different laboratories or under different protocols. An increased incidence of tumours in both sexes of a single species in a well-conducted study, ideally conducted under Good Laboratory Practices, can also provide sufficient evidence. A single study in one species and sex might be considered to provide sufficient evidence of carcinogenicity when malignant neoplasms occur to an unusual degree with regard to incidence, site, type of tumour or age at onset, or when there are strong findings of tumours at multiple sites;*

***limited evidence of carcinogenicity**: the data suggest a carcinogenic effect but are limited for making a definitive evaluation because, e.g. (a) the evidence of carcinogenicity is restricted to a single experiment; (b) there are unresolved questions regarding the adequacy of the design, conduct or interpretation of the studies; (c) the agent increases the incidence only of benign neoplasms or lesions of uncertain neoplastic potential; or (d) the evidence of carcinogenicity is restricted to studies that demonstrate only promoting activity in a narrow range of tissues or organs."*

The inclusion of Good Laboratory Practices would have implications if similar definitions were to be applied to evidence for endocrine disrupting properties as discussed in 4.1.

5.1.1.3 Inclusion of potency considerations

The so-called CMR Working Group, a group of experts from Member States, industry, trade union and Norway, produced a report on the inclusion of potency consideration when setting specific concentration limits for carcinogens[1]. As potency-based cut-offs have also been proposed for endocrine disrupters (section 6), the approach used to derive potency groups for carcinogens is of interest, although they were considered with regards to the classification of intentional mixtures containing classified carcinogens. The report proposed that carcinogens be subdivided into three potency groups of high, medium and low potency, based on a tumorigenic dose descriptor, T25, the dose giving a tumour incidence of 25% in an exposed human population, or in experimental animals after correction for the spontaneous incidence:

- Carcinogens of high potency: T25 value ≤ 1 mg/kg bodyweight/day
- Carcinogens of medium potency: 1 mg/kg bodyweight/day < T25 value ≤ 100 mg/kg bodyweight/day
- Carcinogens of low potency: T25 value > 100 mg/kg bodyweight/day

[1] Commission Working Group on the Classification and Labelling of Dangerous Substances. No date. Guidelines for setting specific concentration limits for carcinogens in annex I of Directive 67/548/EEC - inclusion of potency considerations. [ONLINE. http://ec.europa.eu/environment/chemicals/dansub/pdfs/potency.pdf. ACCESSED 15/12/2011]

The leading principle for deriving the 1 and 100 mg/kg bodyweight/day cut-off values was based on the distribution of another related dose descriptor, TD50, the daily dose rate required to halve the probability of remaining tumorless at the end of a standard life-span, from a database of 492 carcinogens.

5.1.2 REPRODUCTIVE AND DEVELOPMENTAL TOXICANTS

5.1.2.1 Definitions and categorisation

For the purpose of classification, reproductive toxicity is defined as:

"adverse effects on sexual function and fertility in adult males and females, as well as developmental toxicity in the offspring."

The hazard class "Reproductive Toxicity" is differentiated into adverse effects on sexual function and fertility, or on development; and effects on or via lactation.

The legal text also defines "adverse effects" based on the working definitions agreed in IPCS/EHC Document No 225[1]:

Adverse effects on sexual function and fertility

"Any effect of substances that has the potential to interfere with sexual function and fertility. This includes, but is not limited to, alterations to the female and male reproductive system, adverse effects on onset of puberty, gamete production and transport, reproductive cycle normality, sexual behaviour, fertility, parturition, pregnancy outcomes, premature reproductive senescence, or modifications in other functions that are dependent on the integrity of the reproductive systems."

Adverse effects on development of the offspring

"Developmental toxicity includes, in its widest sense, any effect which interferes with normal development of the conceptus, either before or after birth, and resulting from exposure of either parent prior to conception, or exposure of the developing offspring during prenatal development, or postnatally, to the time of sexual maturation. However, it is considered that classification under the heading of developmental toxicity is primarily intended to provide a hazard warning for pregnant women, and for men and women of reproductive capacity. Therefore, for pragmatic purposes of classification, developmental toxicity essentially means adverse effects induced during pregnancy, or as a result of parental exposure. These effects can be manifested at any point in the life span of the organism. The major manifestations of developmental toxicity include (1) death of the developing organism, (2) structural abnormality, (3) altered growth, and (4) functional deficiency."

Reproductive toxicants are classified under two main categories that correspond to those applied to carcinogens; category 1, known or presumed human reproductive toxicants or category 2 suspected human reproductive toxicants. Category 1 is again subdivided into category 1A (known human reproductive toxicant) and 1B (presumed human reproductive toxicant) primarily on the basis of whether the evidence is derived from human or animal studies, respectively. Adverse effects on or

[1] International Programme for Chemical Safety. 2001. Principles For Evaluating Health Risks To Reproduction Associated With Exposure To Chemicals. Environmental Health Criteria No 225. WHO. Geneva, Switzerland.

via lactation are treated separately. The classification scheme and type or strength of evidence associated with each category is reproduced in Table 6.

Classification is made on the basis of an assessment of the weight-of-evidence. The following factors are mentioned as relevant to the evaluation; evidence for substances chemically related to the substance under study, the presence of maternal toxicity in experimental animal studies, relevance of route of administration to humans, relevance of the mode of action for humans.

5.1.2.2 Inclusion of potency considerations

The UN Sub-Committee of Experts on the GHS discussed the scientific issue of the relative potency of reproductive toxicants in mixtures and the OECD recommended that dose cut-off values as a manifestation of relative potency could not be warranted by the current state of scientific knowledge[1].

[1] Organisation for Economic Co-operation and Development .2005. Health hazards, Toxic to reproduction substances, Scientific issue on reproductive toxicity potency. Sub-Committee of Experts on the Globally Harmonized System of Classification and Labelling of Chemicals, Ninth session, 11-13 July 2005, Item 2 (b) (iv) of the provisional agenda.

Table 6. Categorisation of Human Reproductive Toxicants

Categorisation	Definition	Criteria
Category 1	Known or presumed human reproductive toxicant	Substances are classified in Category 1 for reproductive toxicity when they are known to have produced an adverse effect on sexual function and fertility, or on development in humans or when there is evidence from animal studies, possibly supplemented with other information, to provide a strong presumption that the substance has the capacity to interfere with reproduction in humans.
Category 1A	Known human reproductive toxicant	The classification of a substance in Category 1A is largely based on evidence from humans.
Category 1B	Presumed human reproductive toxicant	The classification of a substance in Category 1B is largely based on data from animal studies. Such data shall provide clear evidence of an adverse effect on sexual function and fertility or on development in the absence of other toxic effects, or if occurring together with other toxic effects the adverse effect on reproduction is considered not to be a secondary non-specific consequence of other toxic effects. However, when there is mechanistic information that raises doubt about the relevance of the effect for humans, classification in Category 2 may be more appropriate.
Category 2	Suspected human reproductive toxicant	Substances are classified in Category 2 for reproductive toxicity when there is some evidence from humans or experimental animals, possibly supplemented with other information, of an adverse effect on sexual function and fertility, or on development, and where the evidence is not sufficiently convincing to place the substance in Category 1. If deficiencies in the study make the quality of evidence less convincing, Category 2 could be the more appropriate classification. Such effects shall have been observed in the absence of other toxic effects, or if occurring together with other toxic effects the adverse effect on reproduction is considered not to be a secondary non-specific consequence of the other toxic effects.

5.1.3 SPECIFIC TARGET ORGAN TOXICITY – REPEATED EXPOSURE (STOT-RE)

5.1.3.1 Definitions and categorisation

STOT-RE classification identifies substances that "may present a potential for adverse health effects" and is defined as:

"specific, target organ toxicity arising from a repeated exposure to a substance or mixture. All significant health effects that can impair function, both reversible and irreversible, immediate and/or delayed are included."

Classification under two categories includes the use of cut-off guidance values to discriminate between low and moderate exposure concentrations. The classification scheme is reproduced in Table 7 and cut-off guidance values are discussed further in 5.1.3.2.

Table 7. STOT-RE Categorisation

Categories	**Criteria**
Category 1	Substances that have produced significant toxicity in humans or that, on the basis of evidence from studies in experimental animals, can be presumed to have the potential to produce significant toxicity in humans following repeated exposure. Substances are classified in Category 1 for target organ toxicity (repeat exposure) on the basis of: — reliable and good quality evidence from human cases or epidemiological studies; or — observations from appropriate studies in experimental animals in which significant and/or severe toxic effects, of relevance to human health, were produced at generally low exposure concentrations.
Category 2	Substances that, on the basis of evidence from studies in experimental animals can be presumed to have the potential to be harmful to human health following repeated exposure. Substances are classified in category 2 for target organ toxicity (repeat exposure) on the basis of observations from appropriate studies in experimental animals in which significant toxic effects, of relevance to human health, were produced at generally moderate exposure concentrations. In exceptional cases human evidence can also be used to place a substance in Category 2.

5.1.3.2 Potency-based cut-off guidance values

The CLP regulation provides potency-based cut-off guidance values applicable to toxic effects observed in a 90-day repeated dose study. The guidance values for category 1 and category 2 classifications are reproduced in Table 8 and Table 9, respectively.

Table 8. Guidance values to assist in Category 1 classification

Route of exposure	Units	Guidance values (dose/concentration)
Oral (rat)	mg/kg body weight/day	C ≤ 10
Dermal (rat or rabbit)	mg/kg body weight/day	C ≤ 20
Inhalation (rat)gas	ppmV/6h/day	C ≤ 50
Inhalation (rat) vapour	mg/litre/6h/day	C ≤ 0,2
Inhalation (rat) dust/mist/fume	mg/litre/6h/day	C ≤ 0,02

Table 9. Guidance values to assist in Category 2 classification

Route of exposure	Units	Guidance values (dose/concentration)
Oral (rat)	mg/kg body weight/day	10 < C ≤ 100
Dermal (rat or rabbit)	mg/kg body weight/day	20 < C ≤ 200
Inhalation (rat)gas	ppmV/6h/day	50 < C ≤ 250
Inhalation (rat) vapour	mg/litre/6h/day	0,2 < C ≤ 1,0
Inhalation (rat) dust/mist/fume	mg/litre/6h/day	0,02 < C ≤ 0,2

The CLP states that “the principal argument for proposing such guidance values is that all substances are potentially toxic and there has to be a reasonable dose/concentration above which a degree of toxic effect is acknowledged. Also, repeated-dose studies conducted in experimental animals are designed to produce toxicity at the highest dose in order to optimise the test objective and so most studies will reveal some toxic effect at least at this highest dose”. The parallel argument that has been advanced to apply these guidance values for categorisation of endocrine disrupters is that, as the endocrine system is a communication system, any chemical, if tested at high enough doses would result in an effect on the endocrine system (section 6).

The choice of these values is said to be pragmatic and no scientific basis for the selection of value could be found in the regulatory literature. The UK Committee on Toxicity (COT) in their proposal has

also argued that "*these values (...) have been in place within the framework of the regulatory hazard classification system in Europe since 1967 and are well established and accepted*" (section 6.2). However, the guidance values provided in the Dangerous Substances Directive to classify substances as either toxic or harmful are half the guidance values for the corresponding category 1 and 2 for oral and dermal exposure (see Table 10 and Table 11). No rationale is given for this doubling of the protection threshold.

Table 10. Guidance values to classify a substance as Toxic in the Dangerous Substances Directive (equivalent to STOT-RE Category 1 in the GHS-CLP)

Route of exposure	Units	Guidance values (dose/concentration)
Oral (rat)	mg/kg body weight/day	C ≤ 5
Dermal (rat or rabbit)	mg/kg body weight/day	C ≤ 10
Inhalation (rat)	mg/litre/6h/day	C ≤ 0,025

Table 11. Guidance values to classify a substance as Harmful in the Dangerous Substances Directive (equivalent to STOT-RE Category 2 in the GHS-CLP)

Route of exposure	Units	Guidance values (dose/concentration)
Oral (rat)	mg/kg body weight/day	5 < C ≤ 50
Dermal (rat or rabbit)	mg/kg body weight/day	10 < C ≤ 100
Inhalation (rat) dust/mist/fume	mg/litre/6h/day	0,025 < C ≤ 0,25

5.1.4 ENVIRONMENTAL HAZARDS

There is only one class for environmental hazard under the CLP-GHS scheme, that of substances hazardous to the aquatic environment. This hazard class is further differentiated into acute and chronic aquatic toxicity. The core classification system for substances consists of one acute classification category and three chronic classification categories. The system also introduces a 'safety net' classification (referred to as Chronic Category 4) for use when the data available do not allow classification under the formal criteria but when there are nevertheless some reasons for concern. The classification of a substance into the chronic categories is based on chronic toxicity data when this is available or the combination of acute aquatic toxicity data and environmental fate data (degradability and bioaccumulation data). Classification into the different hazard categories is based on the use of cut-off values applicable to either type of data. The classification categories for substances hazardous to the aquatic environment are reproduced in Table 12.

The reliance on acute toxicity data demonstrates that unless also persistent, bioaccumulative and/or acutely toxic, endocrine disrupters will not be classified as hazardous to the aquatic environment.

Table 12. Classification categories for substances hazardous to the aquatic environment

Acute (short-term) aquatic hazard	
Acute Category 1	96 hr LC50 (for fish) ≤ 1 mg/l and/or 48 hr EC50 (for crustacea) ≤ 1 mg/l and/or 72 or 96 hr ErC50 (for algae or other aquatic plants) ≤ 1 mg/l.
Chronic (long-term) aquatic hazard	
Chronic Category 1	*(i) Non-rapidly degradable substances for which there are adequate chronic toxicity data available* Chronic NOEC or ECx (for fish) ≤ 0,1 mg/l and/or Chronic NOEC or ECx (for crustacea) ≤ 0,1 mg/l and/or Chronic NOEC or ECx (for algae or other aquatic plants) ≤0,1 mg *(ii) Rapidly degradable substances for which there are adequate chronic toxicity data available* Chronic NOEC or ECx (for fish) ≤ 0,01 mg/l and/or Chronic NOEC or ECx (for crustacea) ≤ 0,01 mg/l and/or Chronic NOEC or ECx (for algae or other aquatic plants) ≤ 0,01 mg/l. *(iii) Substances for which adequate chronic toxicity data are not available* 96 hr LC50 (for fish) ≤ 1 mg/l and/or 48 hr EC50 (for crustacea) ≤ 1 mg/l and/or 72 or 96 hr ErC50 (for algae or other aquatic plants) ≤ 1 mg/l and the substance is not rapidly degradable and/or the experimentally determined BCF ≥ 500 (or, if absent, the log Kow ≥ 4)
Chronic Category 2	*(i) Non-rapidly degradable substances for which there are adequate chronic toxicity data available* Chronic NOEC or ECx (for fish) > 0,1 to ≤ 1 mg/l and/or Chronic NOEC or ECx (for crustacea) > 0,1 to ≤ 1 mg/l and/or Chronic NOEC or ECx (for algae or other aquatic plants) > 0,1 to ≤ 1 mg/l. *(ii) Rapidly degradable substances for which there are adequate chronic toxicity data available* Chronic NOEC or ECx (for fish) > 0,01 to ≤ 0,1 mg/l and/or Chronic NOEC or ECx (for crustacea) > 0,01 to ≤ 0,1 mg/l and/or

	Chronic NOEC or ECx (for algae or other aquatic plants) > 0,01 to ≤ 0,1 mg/l *(iii) Substances for which adequate chronic toxicity data are not available* 96 hr LC50 (for fish) > 1 to ≤10 mg/l and/or 48 hr EC50 (for crustacea) > 1 to ≤10 mg/l and/or 72 or 96 hr ErC50 (for algae or other aquatic plants) > 1 to ≤10 mg/l and the substance is not rapidly degradable and/or the experimentally determined BCF ≥ 500 (or, if absent, the log Kow ≥ 4).).
Chronic Category 3	*(ii) Rapidly degradable substances (Note 3) for which there are adequate chronic toxicity data available* Chronic NOEC or ECx (for fish) > 0,1 to ≤ 1 mg/l and/or Chronic NOEC or ECx (for crustacea) > 0,1 to ≤ 1 mg/l and/or Chronic NOEC or ECx (for algae or other aquatic plants) > 0,1 to ≤ 1 mg/l *(iii) Substances for which adequate chronic toxicity data are not available* 96 hr LC50 (for fish) > 10 to ≤ 100 mg/l and/or 48 hr EC50 (for crustacea) > 10 to ≤ 100 mg/l and/or 72 or 96 hr ErC50 (for algae or other aquatic plants) > 10 to ≤ 100 mg/l and the substance is not rapidly degradable and/or the experimentally determined BCF ≥ 500 (or, if absent, the log Kow ≥ 4).
'Safety net' classification	
Chronic Category 4	Cases when data do not allow classification under the above criteria but there are nevertheless some grounds for concern. This includes, for example, poorly soluble substances for which no acute toxicity is recorded at levels up to the water solubility, and which are not rapidly degradable and have an experimentally determined BCF ≥ 500 (or, if absent, a log Kow ≥ 4), indicating a potential to bioaccumulate, will be classified in this category unless other scientific evidence exists showing classification to be unnecessary. Such evidence includes chronic toxicity NOECs > water solubility or > 1 mg/l, or evidence of rapid degradation in the environment.

5.2 REACH

Art. 57(f) of the REACH Regulation allows the identification of substances with endocrine disrupting properties as 'substances of equivalent concern' to Carcinogens, Mutagens and Reproductive and Developmental Toxicants (CMR) or Persistent, Bioaccumulative and Toxic (PBTs) or very Persistent and very Bioaccumulative (vPvBs) substances. For the implementation of the REACH Regulation, ECHA has published Guidance Documents on REACH processes and methods, to be used by industry and authorities. Guidance Documents related to information requirements and chemical safety assessment were reviewed to gain an understanding of the testing requirements likely to apply to substances for which there is little or no existing data. This section will first compare current testing requirements with the tests included in the OECD conceptual framework, before discussing guidance regarding the equivalent level of concern specifically related to endocrine disrupters.

5.2.1 CURRENT TESTING REQUIREMENTS

Article 13.3 states that any new tests should be "conducted in accordance with the test methods laid down in a Commission Regulation or… other international test methods recognised by the Commission or the Agency as being appropriate". The regulation referred to here is the Test Methods Regulation (EC No 440/2008). For some tests, equivalence with a test adopted by the OECD is clearly stated in the Test Methods Regulation, but not for others. ECHA guidance does however tend to refer to OECD test methods.

The testing requirements for substances for registration under REACH are differentiated according to supply tonnage. Testing requirements at a lower supply tonnage generally apply to the higher supply tonnage, unless specific exemptions are clearly stated. The relevant test methods included in the OECD Conceptual Framework are mentioned in relation to testing for repeated dose toxicity (STOT-RE), carcinogens, reproductive toxicants and environmental toxicants. A comparative analysis of the endpoints measured in the peer-reviewed literature, OECD studies included in the Conceptual Framework and for testing under REACH can be found in Annex 3 and specific tables in Annex 3 will be referred to throughout this section. It should be noted that interpretation of the testing requirements will in practice depend on a weight-of-evidence evaluation of existing data and may therefore be different to the minimum requirements as generally interpreted here.

5.2.1.1 STOT-RE

The minimum testing requirements for repeated dose toxicity are summarised in Table 13.

Table 13. STOT-RE testing requirements under REACH by tonnage level

	Tonnage level			
	≥ 1 t/year Annex VII	≥ 10 t/year Annex VIII	≥ 100 t/year Annex IX	≥ 1000 t/year Annex X
Tests required		28 days repeated dose study (TG 407)	90 days repeated dose study (TG 408)	90 days repeated dose study (TG 408)

The default assumption is that oral exposure will be the most representative of human exposure. Different methods can however be applied if dermal exposure or inhalation are thought to be more relevant. **The equivalent methods for short-term (28 days) or sub-chronic (90 days) toxicity for dermal exposure do not include the few endpoints that are relevant to endocrine-mediated toxicity**, neither are they validated for the detection of endocrine disrupters.

Both the short-term and subchronic toxicity studies (TG 407 and TG 408, respectively) are included in level 4 of the OECD Conceptual Framework, however only TG 407 has been validated for the detection of endocrine disrupters (see 3.1). These tests do measure some parameters which are relevant to endocrine-mediated toxicity such the weight and histopathology of the pituitary, adrenals, ovaries and ventral prostate. They do also include endpoints related to neurotoxicity as well as the weight and histopathology of the brain. Some of these endpoints, particularly those related to the thyroid, are optional, and the lack of relevant endpoints is particularly striking for those most relevant to the testicular dysgenesis syndrome (Tables 1-3 in Annex 3). As discussed in section 3.1, the limitations of these studies in terms of screening for endocrine disrupting properties can be related to the fact that only adult animals are exposed and the gross endocrine endpoints themselves lack sensitivity. This point is particularly important for compounds in Annex VIII (tonnage below 100 tons per year) as further tests related to potential carcinogenic or reproductive effects are not generally required.

5.2.1.2 Carcinogens

There are no standard information requirements for substances produced or imported in quantities of less than 1000 tons per year (Annex VII-IX). A carcinogenicity study for substances produced or imported in quantities ≥ 1000 tons per year (Annex X) is only required if the toxicity information already available, together with details of use and human exposure for the substance in question warrant it. The criteria that would trigger a carcinogenicity test for those High Production Volume chemicals (HPV) are the following;

- "*the substance has a widespread dispersive use or there is evidence of frequent or long-term human exposure; and*
- *the substance is classified as mutagen category 3 or there is evidence from the repeated dose study(ies) that the substance is able to induce hyperplasia and/or pre-neoplastic lesions.*
- *If the substance is classified as mutagen category 1 or 2, the default presumption would be that a genotoxic mechanism for carcinogenicity is likely. In these cases, a carcinogenicity test will normally not be required*."

Therefore, some endpoints that are relevant for hormonally mediated cancers are included in the repeated dose toxicity tests (see Tables 19-21 in Annex 3) and may trigger a carcinogenicity study if information on use and human exposure warrant it. The limitations of standard repeated dose studies (TG 407 and TG 408) in terms of the timing of exposure and the sensitivity of the endpoint have already been mentioned above (section 3.1) and again raise serious doubt over the likelihood that potential effects on hormonally mediated carcinogenesis will be detected on the basis of those tests alone.

5.2.1.3 Mutagenicity

Mutagenicity testing requirements are very specifically related to the detection of a potential genetoxic mechanism of action for carcinogenesis or reproductive toxicity. As such, the assays required bear no relevance to the detection of endocrine disrupting properties. Nevertheless, it provides an interesting example of a detailed testing strategy for the detection of a specific mode of action linking the results of specific assays to further testing. The flow chart of the mutagenicity testing strategy given in ECHA guidance is reproduced (Figure 3) to illustrate this point.

It has often been stressed that endocrine disruption is not an endpoint *per se* but a mechanism of action. Further, the potential adverse effects of endocrine disrupters do overlap with carcinogenesis and reproductive toxicity. The application of a similar tiered testing strategy for the detection of endocrine disrupting properties of chemicals is therefore of particular interest and ought to be considered.

The crucial questions here are whether the current testing methods for these classes include the endpoints most sensitive to endocrine disruption and encompass all critical windows of development and potential adverse effects related to endocrine disruption. The strains of animals used in standard carcinogenicity tests are generally poor animal models for hormonally mediated cancers in humans (section 5 in Annex 1, Annex 3 Table 19). While it could be argued that any effect following *in utero* and neonatal exposure could lead to the classification of a chemical substance as a developmental toxicant, it should be noted that some sensitive endpoints related to endocrine disruption (e.g. anogenital distance, nipple retention, neurotoxicity and immunotoxicity) are not routinely measured in a two-generation reproduction study. Further, the prepubertal period has been identified as another critical window of susceptibility to endocrine disruption that is not included in standard reproduction studies (see 5.1.2.1). Further complications arise from the multiplicity of potential mechanisms through which endocrine regulation of development may be disrupted. These issues would need to be addressed, were a similar testing strategy to be devised to detect endocrine disrupting properties.

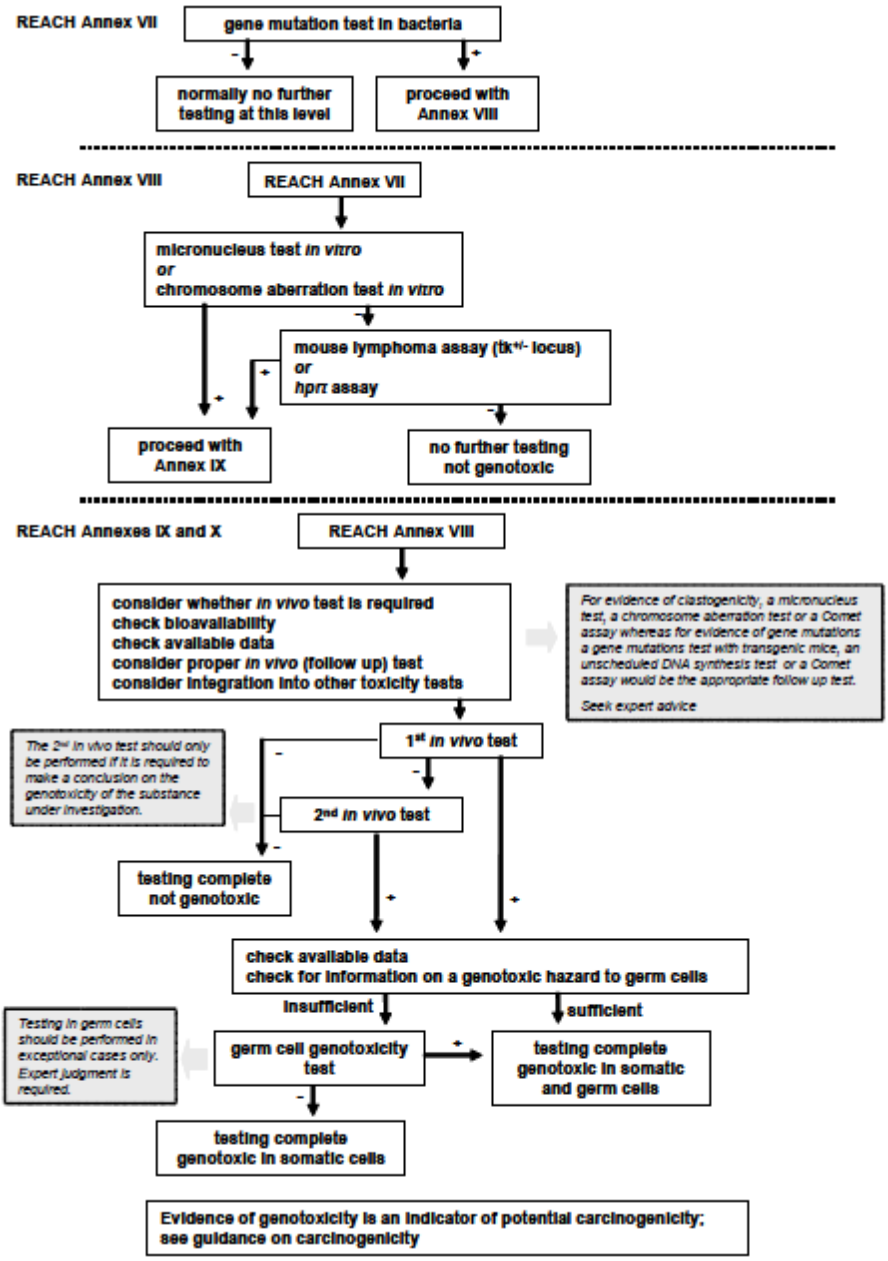

Figure 3. Flow chart of the mutagenicity testing strategy

5.2.1.4 Reproductive and developmental toxicants

The minimum testing requirements for reproductive toxicity are summarised in Table 14.

Table 14. Reproductive toxicity testing requirements under REACH by tonnage level

	Tonnage level			
	≥ 1 t/year Annex VII	≥ 10 t/year Annex VIII	≥ 100 t/year Annex IX	≥ 1000 t/year Annex X
Tests required		Combined repeated dose/reprotoxicity screening test (TGs 421 or 422)	Combined repeated dose/reprotoxicity screening test (TGs 421/422) Prenatal developmental toxicity study (TG 414) in one species, possibly a second species	Prenatal developmental toxicity study (TG 414) in one species, possibly a second species Two-generation reproductive study (TG 416)

Both the combined repeated dose/reprotoxicity screening tests (TG 421/422) and the prenatal developmental study (TG 414) are included as level 4 assays in the OECD Conceptual Framework. The two-generation reproductive toxicity study (TG 416) is included at level 5 of the OECD Conceptual Framework. Neither the combined repeated dose/reprotoxicity screening tests nor the prenatal development study have yet been validated for the detection of endocrine disrupters. Some limited guidance on the interpretation of results from the combined repeated dose/reprotoxicity screening tests was included in the annexes of the OECD guidance document. No such guidance is available for the prenatal developmental study at present. The endpoints monitored in these guideline studies are listed with reference to specific human health endpoints in Tables 3, 6, 9, 12 and 15 in Annex 3.

In the prenatal developmental study (TG 414), animals are exposed from implantation to two days before expected birth and in the combined repeated dose/ reprotoxicity study (TG 421/422), animals are exposed from two weeks prior to mating to four days postnatally. Although these standard tests include exposure during pregnancy, the endpoints related to fertility and gestation maintenance are measured in the parent rather than the subsequent generation. They can therefore not be considered to include exposure during critical windows of development for those endpoints. In the prenatal developmental toxicity study, the aborted pups are checked for gross anomalies. However, differences in terms of developmental milestones between humans and rodents should be borne in mind, as some developmental processes that take place neonatally in the rat take place during the third trimester of human pregnancy. Therefore, although these studies include some endpoints on gestation maintenance and limited developmental effects, they suffer from the same limitation as repeated dose toxicity studies with regards to the exposure period. Further, the lower number of animals used in the combined repeated dose/reprotoxicity study (8-10 parental males and females) will decrease their statistical power. These considerations raise doubts as to the ability

of the current testing requirements to adequately screen for endocrine disrupting properties at tonnage levels below 1000 tons per year (Annex X).

For chemicals with a supply tonnage level in excess of 1000 tons per year, a two-generation study is generally required. In addition, a developmental neurotoxicity study (TG 426) may also be recommended. The OECD Conceptual Framework Guidance document includes the interpretation of the results of a two-generation reproduction study. The deficiencies of the assay in terms of detecting adverse effects of endocrine disrupters compared with the recently validated extended one-generation study (TG 443) have already been briefly discussed in section 3.

5.2.1.5 Environmental toxicants

The minimum testing requirements for environmental toxicity are summarised in Table 15.

Table 15. Environmental toxicity testing requirements under REACH by tonnage level

	Tonnage level			
	≥ 1 t/year Annex VII	≥ 10 t/year Annex VIII	≥ 100 t/year Annex IX	≥ 1000 t/year Annex X
Tests required for pelagic toxicity	Short-term toxicity in invertebrates (daphnia) (TG 202) Growth inhibition study on aquatic plants (TG 201, algae or TG 221, lemna)	Short-term toxicity in invertebrates (daphnia) (TG 202) Growth inhibition study on aquatic plants (TG 201, algae or TG 221, lemna) Short-term toxicity in fish (TG 203)	Short-term toxicity in invertebrates, preferably daphnia (TG 202) Growth inhibition study on aquatic plants (TG 201, algae or TG 221, *lemna*) Short-term toxicity in fish (TG 203) Long-term toxicity on invertebrates (TG 211) Long-term toxicity in fish, either Fish early life stage (TG 210), Fish short-term toxicity on embryo and sac-fry stages (TG 212), Fish juvenile growth test (TG 215)	Short-term toxicity in invertebrates (daphnia) (TG 202) Growth inhibition study on aquatic plants (TG 201, algae or TG 221, lemna) Short-term toxicity in fish (TG 203) Long-term toxicity on invertebrates (TG 211) Long-term toxicity in fish, either Fish early life stage (TG 210), Fish short-term toxicity on embryo and sac-fry stages (TG 212), Fish juvenile growth test (TG 215)
Tests required for toxicity to sediment organisms				Sediment-Water Chironomid Toxicity (TG 218-219)

In addition there are testing requirement for toxicity to sewage treatment plant micro-organisms.

A comparison of the table above and the test methods included in the OECD Conceptual Framework clearly exposes the lack of overlap between current regulatory testing requirements under REACH and methods validated or undergoing validation for the detection of endocrine disrupters. The only test that is common to both schemes is the assay for long-term toxicity in aquatic invertebrates (TG 211) for which no guidance has yet been provided in the Conceptual Framework Guidance Document. ECHA guidance on information requirements clearly states that "*there is no requirement set out in REACH Annexes VII to X to provide information on the endocrine activity of a substance or*

on a substance's reproductive or specific developmental toxicity in aquatic organisms". It does nonetheless provide guidance on the evaluation of existing data for potential endocrine activity of a substance or long-term adverse effects on development and/or reproduction in aquatic organisms in Appendix 7.8-5. It recommends that all available information including non-testing data such as that derived from QSARs and read-across, *in vitro* screening data, and *in vivo* data in vertebrate and invertebrate organism. The recommended integrated assessment of available information is reproduced in Table 16.

Table 16. Integrated assessment of potential endocrine activity in aquatic organisms; based on the evaluation of available information which is not part of the REACH requirements

<table>
<tr><td colspan="3">1) Preliminary indication of potential endocrine activity in aquatic organisms</td></tr>
<tr><td>Estrogen/androgen axis:
- molecular structure
- mammalian toxicity
- in vitro screening</td><td>Thyroid:
- molecular structure
- mammalian toxicity</td><td>Invertebrate systems:
- molecular structure</td></tr>
<tr><td colspan="3">-> determine concern of potential endocrine mode of action of the substance using WoE of all available information, including environmental fate and exposure
-> strong concern may prompt a proposal by the Competent Authority to include the substance in the Community rolling action plan in order to perform a substance evaluation
2) Indication of specific endocrine modes of action in intact aquatic organisms</td></tr>
<tr><td>Estrogen/androgen axis:
- biochemical markers
- morphological changes
(- gonad histopathology)</td><td>Thyroid:
- thyroid histopathology</td><td rowspan="2">Invertebrate systems:
- rare individual cases</td></tr>
<tr><td>Study type:
Fish Screening Assay
Fish Sexual Develpt. Test
Fish Reproduction Test
Fish Full Life-Cycle Test</td><td>Study type:
Amphibian Metamorphosis Assay</td></tr>
<tr><td colspan="3">-> determine concern of potential endocrine mode of action in intact aquatic organisms using WoE of all available information, including environmental fate and exposure
-> strong concern may prompt a proposal by the Competent Authority to include the substance in the Community rolling action plan in order to perform a substance evaluation
3) Characterisation of long-term adverse effects[#]</td></tr>
<tr><td>Estrogen/androgen axis:
- fish (sexual) development
- fish reproduction</td><td>Thyroid:
- amphibian development</td><td>Invertebrate systems:
- development
- reproduction</td></tr>
<tr><td>Study type:
Fish Sexual Develpt. Test
Fish Reproduction Test
Fish Full Life-Cycle Test</td><td>Study type:
Amphibian Metamorphosis Assay</td><td>Study type:
Invertebrate Reproduction or Life-Cycle Tests</td></tr>
<tr><td colspan="3">-> consider use of chronic NOEC for PBT assessment and Chemical Safety Assessment
-> consider classification and labelling according to safety net categories (R52, R53)
-> causal link of adverse effect with an endocrine mode of action may prompt consideration for Annex XV by CA</td></tr>
</table>

5.2.2 EQUIVALENT CONCERN AND ENDOCRINE DISRUPTING PROPERTIES

ECHA has also published a guidance document on the preparation of an Annex XV dossier[1] where it stresses the requirement for "*scientific evidence of probable serious effects to humans or the environment [...] at least equivalent to those that could occur from substances included under points (a) to (e)*" of article 57 of the REACH Regulation.

Therefore equivalent concern can be based either on a comparison of the probable effect with CMRs or with PBT or vPvB criteria. Elements of the document related to equivalency of the effect with CMR and PBT criteria, or more specifically the T-criterion are documented below before summarising guidance on endocrine disrupting properties specifically.

5.2.2.1 T-criterion

An analogy between the potential delayed irreversible effects of endocrine disrupters and persistence and bioaccumulation has been advanced with regards to the uncertainty regarding the effects of endocrine disrupters at low doses (see 4.4.1). ECHA guidance on the subject of the level of concern raised by PBT substances states that:

"*A key part of this concern is that if, ultimately, harmful effects on man or ecosystems are observed, then such effects will be difficult to reverse by control at that stage. Thus, it is the dual potential of uncertainty in being able to say with confidence at what level the substance may be considered safe, along with the serious consequences that could arise from this, that dictate a different approach to risk assessment for these substances*."

In this context, it is useful to briefly examine the T-criterion. The criteria for PBT and vPvBs substances are specified in Annex XIII of the REACH Regulation. The T-criterion is defined as follows:

"*A substance fulfils the toxicity criterion (T-) when:*

– the long-term no-observed effect concentration (NOEC) for marine or freshwater organisms is less than 0,01 mg/l, or

– the substance is classified as carcinogenic (category 1 or 2), mutagenic (category 1 or 2), or toxic for reproduction (category 1, 2, or 3), or

– there is other evidence of chronic toxicity, as identified by the classifications: T, R48, or Xn, R48 according to Directive 67/548/EEC."

These chemical classes refer to classification according to the Dangerous Substance Directive and correspond to carcinogens category 1, mutagen category 1, reproductive toxicant categories 1 and 2 and STOT-RE categories 1 and 2 under the GHS-CLP classification.

[1] European Chemicals Agency. 2007. Guidance for the preparation of an Annex XV dossier on the identification of substances of very high concern. ECHA. Helsinki, Finland.

In terms of other information that may be used to demonstrate a high toxicity potential in relation to equivalent level of concern in a weight of evidence approach, endocrine disrupting effects are specifically mentioned;

"*Substantial evidence of long-term adverse effects. Such evidence may include endocrine disrupting effects for example [...] in cases where they cannot be assessed with conventional hazard assessment methodology. This should be considered on a case-by-case basis.*"

The main point to note here is that following the reasoning for an analogy between PBTs and endocrine disrupters, the T-criterion cut-offs are lower with respect to reproductive toxicity or STOT-RE than those currently included in the various proposals (see section 6). It should nonetheless be borne in mind that all criteria for persistence, bioaccumulation and toxicity have to be met for a substance to be designated as a PBT chemical.

5.2.2.2 CMRs

In contrast with the T-criterion, equivalency of concern is based on CMR category 1. A number of examples of serious effects that would give rise to equivalent concern are given and those are consistent with endpoints measured in STOT-RE studies. No reference is made to potency-based cut-off values for equivalent concern.

5.2.2.3 Endocrine disrupting properties

ECHA guidance on the preparation of an Annex XV dossier argues for a weight-of-evidence approach to the assessment of equivalent concern on the basis of endocrine disrupting properties;

"*Given the complexities of the possible mechanisms and effects of endocrine active substances it is unlikely that the results from isolated screening assays will be sufficient to confirm that the substance has potential to cause endocrine disrupting effects in humans or wildlife. Therefore a weight of evidence approach is needed*."

This line of reasoning appears to be at odds with the application of potency-based cut-off values to differentiate endocrine disrupters of high regulatory concern.

The following criteria were proposed when evaluating the available data;

- "*The balance of positive and negative responses observed in both in vitro and in vivo assays.*
- *The nature and range of the biological effects observed in assays intended to identify and characterize hazards. Again, does a potential estrogenic substance lead to the observation of a pattern of consistent changes in estrogen related target tissues and developmental endpoints*
- *The shape of the dose-response curves when available. For example, does the dose response in the in vivo screen for an endocrine mechanism show a correspondence to the dose response of those endocrine related changes?*
- *The severity and magnitude of effects induced. For example, potent estrogens halt the estrous cycle and potent antiandrogens cause malformations of the male reproductive tract. Both are examples where fertility and reproduction are impaired*

- *The presence or absence of responses in multiple taxa. Endocrine systems are conserved, and a substance acting via an endocrine mechanism is expected to cause endocrine related effects in most or all taxa having those endocrine pathways. For example, estrogenic, androgenic, and thyroid effects would be typically expected to occur in several vertebrate classes for a substance truly acting via a relevant endocrine mechanism*"

5.3 PLANT PROTECTION PRODUCT REGULATION

The new Plant Protection Product Regulation contains a proposal for "cut-off" criteria for the approval of active substances, safeners and synergists based on hazard properties of the substance. Any substance meeting the following "cut-off" criteria based on the hazard properties of the active substances (Annex II 3.6-3.7) shall no longer be approved:

- Mutagen category 1
- Carcinogen or toxic for reproduction category 1, unless the exposure to human is negligible
- Endocrine disrupters, unless the exposure to human is negligible
- POPs (persistent organic pollutant)
- PBT (persistent bioaccumulating and toxic)
- vPvB (very persistent and very bioaccumulating)

With respect to endocrine disrupters specifically, Annex II 3.6.5. states that:

"*An active substance, safener or synergist shall only be approved if, (…) it is not considered to have endocrine disrupting properties that may cause adverse effect in humans, unless the exposure of humans (…), under realistic proposed conditions of use, is negligible, (…).*"

Annex II 3.8.2. relates to non-target organisms:

"*An active substance, safener or synergist shall only be approved if, (…) it is not considered to have endocrine disrupting properties that may cause adverse effects on non-target organisms unless the exposure of non-target organisms to that active substance in a plant protection product under realistic proposed conditions of use is negligible*."

In contrast to the Pesticides Directive 91/414/EEC, detailed data requirements for active substances and products are no longer included in the annexes of the new Regulation (EC) 1107/2009. They are detailed in separate legal instruments (Commission Regulations on data requirements for active substances (544/2011) and products (545/2011)). These data requirements were reviewed to identify any testing method that is also included in the OECD Conceptual Framework and assess the likelihood that endocrine disrupting properties will be adequately detected under the testing framework for approval.

5.3.1 CURRENT TESTING REQUIREMENTS

5.3.1.1 Human health

The following test methods listed in the Regulations on data requirements under toxicological and metabolism studies are included in the OECD Conceptual Framework;

- Oral 28-day study (TG 407)
- Oral 90-day study (TG 408) for both dog and rat
- Combined chronic toxicity/carcinogenicity test (TG 451-3)
- Two- generation reproduction toxicity study (TG 416)
- Developmental toxicity study (TG 414)

With regards to reproductive health, a two-generation study is always required. This assay measures many endpoints relevant for endocrine disruption (see tables 3, 6, 9, 12, 15, 18 and 21 in Annex 3), and animals are exposed *in utero* and neonatally. Interpretation of the results of a two-generation reproduction study in terms of the detection of endocrine disrupters is detailed in the OECD Conceptual Framework Guidance document. The deficiencies of the assay in terms of detecting adverse effects of endocrine disrupters as well as a comparison with the recently validated extended one-generation study (TG 443) have already been discussed in section 3.1.1.

A few endpoints such as the histology of breast, prostate or testis have been argued to be able to indicate a carcinogenic potential for a substance in those target organs and are included in the repeated dose toxicity tests (TG 408) and combined chronic toxicity/carcinogenicity study (TG 451-3) (see table 21 in Annex 3). Limited guidance on the interpretation of results from these tests is included in the OECD Conceptual Framework guidance document, however such guidance is generally more adequate for reproductive toxicity than carcinogenic potential. Despite the long period of exposure in the combined chronic toxicity/carcinogenicity study, young healthy animals are dosed and the method does therefore not include exposure during critical stages of development. Further, the adequacy of animal strains generally used in such standard tests to detect hormonally-mediated cancers has been questioned (Section 5 in Annex 1 and Table 19 in Annex 3). These issues raise serious doubt over the likelihood that effects related to hormonally mediated carcinogenesis can be detected on the basis of those tests alone.

The repeated dose 90-day study (TG 408) includes endpoints of a neurobehavioral functional screening battery (Table 24 in Annex 3). However, the exposure period does not include critical stages of development and this test is not designed to detect developmental neurotoxicity. No endpoints related to the metabolic syndrome or developmental immunotoxicity are measured under current testing requirements for PPPR (Table 27 in Annex 3).

In spite of those omissions, provided that endocrine sensitive endpoints are included, it may be argued that a two-generation reproduction study (TG 416) will detect endocrine disruption effects modulated via the same mode(s) of action that may be involved in some of the adverse effects not currently included in method. It should however again be stressed that as evidenced by the results of both the scientific summary (Annex 1) and the Detailed Review Paper on Novel Endpoints and

Assays[1], the adverse effects of endocrine disrupters are not limited to those of the disruption of the estrogen, androgen, thyroid and steroidogenesis pathways and even level 5 assays in the OECD Conceptual Framework have limitations in terms of detecting adverse effects. These points echoe those of the scientific report of the EFSA Task Force on Endocrine Active Substances[2]:

"*However, it is also evident that existing standard protocols may not cover all potential effects that could be induced by EAS. For example, effects on gastro-intestinal hormones, or reproductive senescence are not covered. Similarly, while existing protocols may allow for the measurement of a number of endogenous hormones, such measurements are not always undertaken, or the results may be confounded by stress or the inconsistency of diurnal variation*s [...]. *It is because of the inconsistency in some of these measurements, revealed by inter-laboratory validation studies, that they are not obligatory in standard test protocols*."

The EFSA EAS Task Force report also highlights the potential value of *in vitro* screening assays for the hazard identification of EDCs:

"*Such assays will be useful for screening substances suspected of having endocrine activity, or structurally related to known EAS, and for providing mechanistic information, but at present are not required routinely for substances deliberately added to food or feed.*"

The EFSA Panel on Plant Protection Products and their Residues reviewed its Opinions issued in 2006 and 2007 related to the revision of Annexes II and III to Council Directive 91/414/EEC (data requirements) concerning the placing of plant protection products on the market[3]. It stressed that additional data requirements would be required when the new Regulation is in place:

""*It will also be necessary to define additional data requirements to address certain issues e.g. certain pesticides which may disrupt the endocrine system, safeners, co-formulants and synergists, once the new Regulation on the placing of plant protection products on the market is in place."*

5.3.1.2 Wildlife

The following test methods listed in the Regulations on data requirements under ecotoxicological studies are included in the OECD Conceptual Framework;

- Bird subchronic toxicity and reproduction (TG 206)
- Fish life cycle test (only if triggered by bioconcentration factor or degradability studies)

Many other tests including field studies, mesocosm studies or effects on earthworms, bees, plants and micro-organisms are included in the Regulation on data requirements and this diminutive 'list' illustrates the very poor match with the OECD Conceptual Framework. One reason for this almost complete lack of overlap between the two testing schemes could be that regulatory data

[1] RTI International. 2011. Draft Detailed Review Paper State of the Science on Novel In Vitro and In Vivo Screening and Testing Methods and Endpoints for Evaluating Endocrine Disruptors. OECD. Paris, France.

[2] European Food Standard Agency. 2010. Scientific report of the Endocrine Active Substances Task Force. EFSA Journal. 8(11):1932. [59 pp.] doi:10.2903/j.efsa.2010.1932. Available online: www.efsa.europa.eu/efsajournal.htm

[3] Panel on Plant Protection Products and their Residues. 2009. Updating the opinion related to the revision of Annexes II and III to Council Directive 91/414/EEC concerning the placing of plant protection products on the market – Toxicological and metabolism studies. EFSA Journal. 1166:1-7

requirements rely on short-term studies rather than the multigenerational studies necessary to demonstrate **adverse effects** of endocrine disrupters on wildlife **populations**. Due to their shorter life-span and lesser ethical concerns over their welfare, many multigenerational tests in invertebrate taxa for example are relatively inexpensive compared to their mammalian equivalents. Regardless of the mismatch in terms of the length of the studies, some taxa such as amphibians are omitted altogether. There may be scope to detect endocrine disrupting effects in some of the tests required that are not included in the OECD Conceptual Framework, however no guidance on the interpretation of the results is available in this context. Under the current testing requirements as set out in the Regulations on data requirements for active substances (544/2011) and products (545/2011), there appears to be only scope to detect endocrine disrupting effects on birds, arguably fish and mammals from data generated from mammalian assays, but not in other taxa (see tables 30, 33, 36, 39, 42, 45, 48 in Annex 3).

5.3.2 HAZARD BASED CUT OFF CRITERIA

The introduction of hazard based "cut-off" criteria has polarised opinions. On the one hand, it is argued that regulating on the basis of hazard rather than risk assessment will lead to restrictions on the use of products that are relatively safe and that may have beneficial effects in terms of food security. An impact assessment of the "cut-off" criteria carried out by the United Kingdom Pesticides Safety Directorate estimated that 5-15% of active substances may no longer be approved[1]. On the other hand, the additional protection afforded in terms of potential health benefits has been estimated to have an upper bound range of €3,568 to €7,160 billion over the coming 30 years for the maximum exposed farm worker population[2]. It should however be noted that without defined criteria for the assessment of endocrine disrupters, these figures remain purely speculative.

The rationale for applying hazard based cut-off criteria generally and in the context of endocrine disrupters particularly can be related to scientific uncertainty and the resultant difficulties in deriving an exposure dose that may be considered safe. In other words, a risk assessment cannot be carried with a satisfactory level of certainty. This is directly related to the issues of low dose, irreversibility and exposure during critical stages of development discussed in sections 4.3 and 4.4.

It should be noted that these cut-off criteria are not applied in the legislative text with complete disregard to potential exposure or consideration for the balance of potential risks and benefits. Article 7 of the PPPR allows for derogation from these cut-off criteria (with the exclusion of human carcinogen category 1A and human reproductive toxicant category 1A):

"*By way of derogation from paragraph 1, where on the basis of documented evidence included in the application an active substance is necessary to control a serious danger to plant health which cannot be contained by other available means including non-chemical methods, such active substance may be approved for a limited period necessary to control that serious danger but not exceeding five years even if it does not satisfy the criteria set out in points 3.6.3, 3.6.4, 3.6.5 or 3.8.2 of Annex II,*

[1] Pesticides Safety Directorate. 2008. Assessment of the impact on crop protection in the UK of the 'cut-off criteria' and substitution provisions in the proposed Regulation of the European Parliament and of the Council concerning the placing of plant protection products in the market. PSD. York, United Kingdom.

[2] Policy Department Economic and Scientific Policy. 2008. The benefits of strict cut-off criteria on human health in relation to the proposal for a Regulation concerning plant protection products. European Parliament's Committee on the Environment, Public Health and Food Safety. Brussels, Belgium.

provided that the use of the active substance is subject to risk mitigation measures to ensure that exposure of humans and the environment is minimised."

These criteria are also not applicable if exposure is expected to be negligible. This negligible exposure criterion has been defined in terms of dietary exposure to residues as less than 0.01 mg/kg and the basis for the selection of this value is not specified. Further, negligible exposure in other situations such as occupational exposure of farmers as well as that of their family, or bystanders in rural area remain to be defined.

The considerations above highlight that it is difficult to separate hazard or risk assessment from risk management measures. Ultimately, what constitutes an acceptable risk is not a scientific but a socio-political decision and this may differ depending on the specific use of a substance. It would indicate that it should be possible to devise decision criteria for the assessment of endocrine disrupters that are common to different regulatory "silos" in terms of hazard identification but approaches to risk management whether hazard or risk based may differ depending on the proposed use for the substance.

5.4 BIOCIDAL PRODUCT REGULATION

The compromise text resulting from the informal trilogues was formally agreed by the European Parliament on 19th January 2012. It is expected that the Council will also approve this text during Spring. Article 5 of the latest text details exclusion criteria that would preclude the granting of authorisation for a given active substance. These criteria are essentially equivalent to the hazard cut-off criteria referred to above in the context of the PPPR. Carcinogens, mutagens or reproductive toxicants category 1, PBT, vPvB, and **substances considered as having endocrine disrupting properties that may cause adverse effects in humans (according to the criteria to be developed by the Commission by 13 December 2013) OR which are identified in accordance with Articles 57(f) and 59 (1) of REACH) shall not be approved**, unless the risk to humans is negligible.

Annex II sets out the information required for the preparation of an authorisation dossier. Two types of information are listed; compulsory data that form the basis of a core dataset, and additional data that may be required if triggered by 'alerts' identified in the core dataset. The relevant core information requirements and the equivalent OECD guideline study are listed in Table 17.

Table 17. Relevant core dataset required for repeated dose toxicity, reproductive toxicity, carcinogenicity and ecotoxicological studies

Repeated Dose toxicity	Short-term 28-day repeated dose toxicity study (TG 407) Sub-chronic 90-day repeated dose toxicity study (TG 408)
Reproductive toxicity	Prenatal developmental toxicity study in the rabbit (TG 414) Two-generation reproductive toxicity study (TG 416) or the extended one generation reproductive toxicity study.
Carcinogenicity	Combined carcinogenicity and long-term repeated dose (TG 451-3) *Carcinogenicity in a second species*
Toxicity to aquatic organisms	Short-term toxicity testing on fish (TG 203) Short-term toxicity testing in Daphnia (TG 202)

Upon evaluation of the core dataset, additional studies may be required. Paragraph 8.13.3 makes specific mention of endocrine disruption and states that "if there is any evidence from *in vitro*, repeat dose or reproduction toxicity studies, that the active substance may have endocrine disrupting properties then additional information or specific studies shall be required:

- To elucidate the mode/mechanism of action
- Provide sufficient evidence for relevant adverse effects"

Furthermore, neurotoxicity including developmental neurotoxicity and immunotoxicity including developmental immunotoxicity may also be required.

Paragraph 9.10 also refers to the identification of endocrine activity with respect to ecotoxicological studies, although no specific test or alerts are mentioned. Other relevant ecotoxicological tests of relevance to the assessment of endocrine disrupters are indicated if long-term exposure is expected and listed in table).

Table 18. Relevant additional ecotoxicological tests

Long-term toxicity on fish	Fish early life stage test Fish short-term toxicity on embryo and sac-fry stages Fish juvenile growth test Fish full life cycle test
Long-term toxicity testing on invertebrates	Daphnia growth and reproduction study Other species reproduction and growth (e.g. Mysid) Other species development and emergence (e.g. Chironomus)
Studies on sediment dwelling organisms	
Effects on birds	Effects on reproduction
Effects on arthropods	Effects on honey bees Other non-target terrestrial arthropods, e.g. predators

The minimum information requirements are more or less equivalent to those of PPPR and high production volume chemicals under REACH, at least for human health, and the same limitations of those batteries of tests will apply. The inclusion of more detailed additional data requirements recognises and attempts to address some of these shortcomings. The explicit inclusion of developmental neurotoxicity and immunotoxicity are examples of this. However, a carcinogenicity study that includes exposure during critical life stages or the multigeneration ecotoxicological studies incorporated in level 5 of the OECD Conceptual Framework are still absent. Further, how these additional requirements are implemented in practice remains to be seen. Concerns over the ability to detect endocrine disrupters on the basis of 'alerts' in short-term adult assays are discussed further in the following section.

5.5 SOME CONCLUDING REMARKS ON TESTING REQUIREMENTS

From an ecotoxicology standpoint, the lack of standard tests for many sensitive endpoints and endocrine modes of action is widely acknowledged and recognised. From a human health perspective, this has been much more controversial and it has been contended that many of the endpoints in endocrine organs traditionally measured in guideline studies would raise toxicological alerts as to the endocrine disrupting potential of substances, thereby triggering additional testing. The information collated on endpoints in Annex 3 of this report is summarised in two separate tables in this section. Table 19 lists the endocrine relevant endpoints that would be routinely measured according to the current test requirements whilst Table 20 lists endpoints that have been suggested either in the peer-reviewed literature surveyed for the scientific summary (Annex 1) or suggested in the OECD Draft Detailed Review Paper on Novel Endpoints and Assays[1].

[1] RTI International. 2011. Draft Detailed Review Paper State of the Science on Novel In Vitro and In Vivo Screening and Testing Methods and Endpoints for Evaluating Endocrine Disruptors. OECD. Paris, France.

Table 19. Endocrine relevant endpoints measured as part of current testing requirements

Endocrine relevant endpoints	OECD Test methods	Comments
Malformation of external genitalia	TG 414, TG 416	
Malformation of internal sex organs (epipidymes, testes)	TG 416	
Poor semen quality	TG 416	
Changes in serum hormones	Optional in TG 407	Adult exposure only
Age at first estrus	TG 416	
Mammary gland development	TG 408, TG 451-3	Adult exposure only
Weight or pituitary and adrenals	TG 407, TG 408, TG 451-3, TG 416	
Uterine weight	TG 408, TG 451-3, TG 416	
Histopathology of ovaries	TG 407, TG 451-3, TG 416	
Weight of ovaries	TG 408, TG 451-3, TG 416	
Estrous cyclicity	TG 416	
Precoital interval	TG 421-422, TG 416	
Number of implantations, corpora lutea, pre and post implantation loss	TG 421-422, TG 414, TG 416	
Histopathology of uterus	TG 407, TG 408, TG 421-422, TG 451-3, TG 416	
Duration of gestation	TG 414, TG 416	
Dystocia	TG 421-422, TG 416	
Placental weight	TG 416	
Number, sex and weight of pups	TG 421-422, TG 414, TG 416	
Number of live and still births	TG 421-422, TG 414, TG 416	
Signs of gross anomalies	TG 414, TG 416	
Weight of ventral prostate	TG 407, TG 408, TG 416	
Prostate histology	TG 407, TG 408, TG 421-422, TG 451-3, TG 416	
Leydig cell nodules/hyperplasia	TG 407, TG 408, TG 451-3	Adult exposure only
Thyroid histopathology	TG 408, TG 421-422, TG 451-3, TG 416	
Serum T3, T4 and TSH	Optional in TG 407 and TG 408	Adult exposure only
Thyroid weight	TG 408, TG 451-3, TG 416	
Brain weight	TG 407, TG 408	Adult exposure only
Brain histopathology	TG 407, TG 408	Adult exposure only
Neurobehaviour battery	TG 407, TG 408	Adult exposure only

TG 407: Short-term 28-day repeated dose toxicity study, TG 408: Sub-chronic 90-day repeated dose toxicity study, TG 414: Prenatal toxicity study, TG 416: Two-generation reproductive toxicity study, TG 421-422: Combined reproduction and 28-day repeated dose study, TG 451-3: Combined carcinogenicity and long-term repeated dose study

Table 20. Endocrine relevant endpoints not measured as part of current testing requirements

Endocrine relevant endpoints	Validated test methods	Comments
Anogenital distance, retained nipples	TG 443	
Retinoid levels, plasma levels or hepatic expression of IGF1		Suggested in the Detailed Review Paper
Age at vaginal opening	Female pubertal assay	
Stress response - corticosterone		Suggested in the Detailed Review Paper
Number of ovarian follicles	TG 443	
Reproductive senescence		
Marker genes: Calbindin-D9k, Hox genes and gap junction protein connexion 26 and 43		Proposed to be more sensitive than uterine weight
Serum vitamin D, vitamin D responsive genes		Suggested in the Detailed Review Paper
Bone histology		Suggested in the Detailed Review Paper
Peroxisome proliferation		Suggested in the Detailed Review Paper
Sexually dimorphic nucleus of the preoptic area (SDN-POA), the anteroventral periventricular nucleus (AVPV) and the locus coeruleus.		
Neurotoxicity module	TG 443	
Thyroid disruption – additional endpoints		Suggested in the Detailed Review Paper
Immunotoxicity module	TG 443	

TG 443: Extended one-generation reproduction toxicity study

The bone of contention lies with the sensitivity and specificity of the endpoints routinely measured. It should be clear that traditional endocrine endpoints focus primarily on the weight and histopathology of endocrine organs, in addition to reproductive endpoints that are not specific to endocrine disruption. Novel endpoints suggested in the OECD draft detailed review paper and in the peer-reviewed literature tend to address evidence for a specific endocrine disruption mode-of-action, but is argued by some as falling short of evidence for an adverse effect. Notwithstanding the limitations of a two-generation reproductive toxicity assay as recognised by the OECD EDTA (see 3.1.1), weight and histopathology of endocrine organs are regarded as relatively insensitive endpoints. This problem is further compounded when these endpoint are supposed to raise toxicological alerts for irreversible developmental effects. This is best illustrated by the data used for the perchlorate case study carried out to support the guidance document for the OECD Conceptual Framework[1]. The available data used for the case study is summarised in Table 21 and Table 22 showing data from standard assays and non-standard assays considered similar to guideline studies, respectively. Effects on thyroid weight, histopathology as well as thyroid hormone levels were

[1] Organisation for Economic Co-operation and Development. 2011. Guidance Document (GD) on Standardised Test Guidelines for Evaluating Chemicals for Endocrine Disruption (No. 150).Case Studies using example chemicals (Perchlorate). Draft v1. Document n° ENV/JM/TG/EDTA(2011)14. EDTA. Paris, France

detected in a two-generation reproductive toxicity assay in all generations. No short-term or subchronic repeated dose toxicity study was available. A non-standard assay where animals had been exposed for 90-days detected effects on the weight and histopathology of the thyroid; however these effects were reversible after a recovery period of 30 days. The measurement of thyroid hormones is optional in a subchronic repeated dose study but was measured in this non-standard study. Changes in TSH and thyroid hormones were observed at all dose levels, therefore at levels below effects on thyroid weight and histopathology. Two non-standard studies considered to be similar to a combined chronic toxicity/carcinogenicity study were also included in the case study dataset. However, the doses administered were extremely high and therefore the results cannot be used to comment on the ability of this test to detect thyroid-disrupting effects. Nor would alerts have been raised from ecotoxicological studies under the current testing regime. It should therefore be clear that on the basis of current testing requirements, there is a real potential for missing or disregarding endocrine-disrupting effects.

The problem can be illustrated further by considering the outcome of attempts to anticipate the endocrine disrupting potential of pesticides on the basis of the data provided by current testing and information requirements. The UK Chemicals Regulation Directorate conducted a case study involving 15 pesticides, with the aim of assessing their endocrine disrupting potential. It was concluded that

"…in many instances the level of detail in end-point tables and EFSA conclusions was inadequate to permit completion of the evaluation for ED effects / relevance. Even going back to Draft Assessment Reports (DARs) in many cases did not provide sufficient detail, particularly on the severity of effects"[1].

Taken together, this highlights the fact that testing requirements need to be adapted to adequately allow the detection of EDCs by requesting different tests, in mammals or other species, and/or adding validated endpoints to existing tests. Detailing such a testing battery is beyond the scope of this report and currently hindered by the lack of data available from standard guideline studies included in the OECD Conceptual Framework.

[1] CRD 2011, Evaluation of a selection of pesticides against the proposed ED criteria and flow chart, kindly provided by Dr Susie Brescia, CRD

Table 21. Test guideline studies used for the perchlorate case study

Wildlife screens and tests	
OECD TG 229: Fish short-term reproduction assay	There were no effects on **fecundity**, **hatching**, **male histology** or **VTG**. There was an increase in the numbers of **atretic ovarian follicles** and a decrease in the number of **stage 1A oocytes** (LOEC = 5.6 mg/l), but it is unclear whether this effect was related to endocrine disruption.
Fish Lifecycle Toxicity Test (USEPA OPPTS 850.1500)	There were no significant changes in **any apical endpoint,** including reproduction and growth, at any concentration in either the F0 or F1 generations (NOEC > 1090 µg/l). However, female **thyroid follicular hypertrophy** (prevalence: 3 out of 12 fish) and **thyroid colloid depletion** (6 out of 14 fish) occurred, with an overall NOEC of 273 µg/l in the F0 fish.
OECD TG 231: Amphibian Metamorphosis Assay (AMA)	**Whole-body length**, **snout-vent length** and **wet wt**. all increased in *X. laevis* larvae (by about 5-20%) in response to perchlorate, with NOEC values generally in the range 62.5-125 µg/l. Hind limb length decreased by 20-30% (NOEC = 125 µg/l). **Thyroid histology** was generally altered at all concentrations (NOEC <62.5-62.5 µg/l), including increases in thyroid epithelial cell height and thyroid volume/area.
OECD TG 206: Avian Reproduction Test	There were no significant effects on most endpoints, but **eggshell thickness** was reduced to 91% of the control value at the LOEC (NOEC = 100 mg/kg feed), and **male weight gain** was reduced to 76% of the control value at the LOEC (NOEC = 500 mg/kg feed).
Mammalian screens and tests	
Peripubertal Male Rats Assay (US EPA OPPTS 890.1500)	**T4** was decreased in a dose-dependent manner from 125 mg/kg/day, **TSH** was increased at the same doses, whilst T3 was unaffected. **Thyroid histology** was significantly altered at all doses with a clear dose-dependent decrease in colloid area and increase in follicular cell height. No effects on preputial separation or reproductive tissues weight were observed.
Peripubertal Female Rats Assay (US EPA OPPTS 890.1450)	**T3 and T4** were decreased in a dose-dependent manner from 125 mg/kg/day, **TSH** was increased at 250 and 500 mg/kg. Thyroid histology was significantly altered at all doses with a dose-dependent decrease in colloid area and increase in follicular cell height. No effects on vaginal opening or weights of uterus and ovaries were observed. There were no treatment-related effects on estrous cyclicity.
OECD TG 416: Two-Generation Reproduction Toxicity Study	There were **no effects** on **reproduction or gestation length and no deaths or abortion**s. There were **no effects on mating, fertility, sperm parameters or estrus cyclicity** in either P or F1 generations. There were **no changes in numbers of live pups, viability, sex ratios or pup body weights**. Endpoints of sexual maturity (e.g. VO, PPS) or AGD were not determined. Dose-dependent changes in **thyroid weight** and **histopathology and hormone levels** were observed in all generations. Relative thyroid weights were increased at 30 mg/kg/day (♀) and 3 & 30 mg/kg/day (♂) in the P generation, whilst in the F1 generation thyroid weights were increased at all doses (♀) and 3 & 30 mg/kg/day (♂). Hypertrophy and hyperplasia of the thyroid follicular epithelium increased in incidence and severity in a dose related manner. Reduced T4 & T3 and increased TSH also occurred in a dose-related manner but were not consistent across the generations and sexes. Two male rats from the F1 30 mg/kg/day group developed thyroid adenomas, compared to none in controls. These animals had had 19 weeks of dosing.

Table 22. Non standard assays - are those where the method described is considered to be scientifically very similar to that given in the appropriate Test Guideline but with significant methodological differences

Wildlife screens and tests	
Similar to	
OECD TG 229: Fish Short Term Reproduction Assay	Adult zebrafish (*Danio rerio*) were exposed to measured perchlorate concentrations of 18 and 677 mg/l for up to **8 weeks**. Both concentrations caused **thyroidal angiogenesis**, and the lower concentration also caused **thyroidal hyperplasia** and **colloid depletion**. However, reproductive performance (**fecundity**) was only depressed (by ~85%) at the higher concentration (NOEC = 18 mg/l).
OECD TG 234: Fish Sexual Development Test	Larval zebrafish (*Danio rerio*) were exposed from 3 d post-fertilisation for 30 d to nominal Perchlorate concentrations of 100 or 250 mg/l. Both treatments caused **hypothyroidism** which could be reversed by co-treatment with exogenous T4. Perchlorate also suppressed **growth** (body length) (by ~15% at the LOEC) and skewed the **sex ratio** towards females (from 48% in controls, to 58% at the LOEC, and to 65% at the top concentration), an effect which was also counteracted by exogenous T4. The NOEC for these effects was < 100 mg/l.
Fish Lifecycle Toxicity Test (FLCTT) (USEPA OPPTS 850.1500)	Three-spined sticklebacks (*Gasterosteus aculeatus*) were exposed for up to 1 year, through the stage of reproductive activity, to measured concentrations of perchlorate ranging from 30 to 100 mg/l. There were **no observed changes** in **reproductive behaviour** or **reproductive output** of adults exposed for 22 d to concentrations up to 18.6 mg/l. However, in fish exposed throughout development, **swimming performance**, **reproductive behaviour**, **survivorship** and **recruitment** were affected at all concentrations (NOEC < 30 mg/l), and only 25% of males produced fry at the LOEC (30 mg/l). Reproductive activity was almost completely inhibited at 100 mg/l, and no males produced fry at this concentration.
OECD TG 231: Amphibian Metamorphosis Assay	Clawed frog (*Xenopus laevis*) eggs and larvae were exposed to measured perchlorate concentrations between 5 and 425,000 µg/l for 70 d. **Forelimb emergence** was reduced (by ~50% at the LOEC) at all concentrations (NOEC < 5 µg/l), while NOECs for reduced **hind limb length** and reduced **tail resorption** were both 5 µg/l. In a USEPA standard screening test, in which tail resorption in stage 60 larvae was measured, 19,800 µg perchlorate /l exposure for 14 d reduced **tail resorption** from 96 to 17%.
OECD TG 231: Amphibian Metamorphosis Assay	Clawed frog (*Xenopus laevis*) larvae were exposed to a measured concentration of perchlorate at 20 mg/l for 12 d. This treatment **delayed development**, and **hind limb length** was reduced by ~50%. These changes were accompanied by **thyroid histopathology**, reduced expression of genes regulated by thyroid hormone, and up-regulation of thyroid-stimulating hormone tshb-A mRNA.
OECD TG 231: Amphibian Metamorphosis Assay	Clawed frog larvae (*Xenopus laevis*) at stage 51-54 were exposed to measured perchlorate at 16, 63, 250, 1000 and 4000 µg/l for 14 d. **Metamorphosis** was significantly retarded (NOEC = 63 µg/l), but **histological effects** were observed at the lowest concentration (NOEC < 16 µg/l). A similar experiment with stage 51 larvae exposed to 8, 16, 32, 63 and 125 µg/l gave a NOEC for **retarded metamorphosis** of 63 µg/l, and a NOEC for **thyroid hypertrophy** of 32 µg/l.
Larval Amphibian Growth and Development Assay (LAGDA)	*X. tropicalis* larvae (<48 h post-hatch) were exposed for up to 40 weeks to measured perchlorate concentrations of 56, 167, 500 and 1500 µg/l. **None of these concentrations caused significant effects on metamorphosis, body size of adults, or gonadosomatic indices of adults.** However, some **thyroid histopathology** was observed (NOEC = 170 µg/l), and **vitellogenin** levels in females were increased by a factor of approximately 1.7 (NOEC = 56 µg/l).

OECD TG 206: Avian Reproduction Test	Laying female Japanese quails (*Coturnix japonica*) were exposed to perchlorate at 2000 or 4000 mg/l in drinking water. Thyroid status was examined in the resulting **embryos**. **Hypothyroidism** was observed in the adults of both treatment groups (NOEC < 2000 mg/l), and **egg production** was decreased in the high dosage group only (NOEC = 2000 mg/l). The embryos from both groups also experienced **hypothyroidism**, as evidenced by **thyroid gland hypertrophy** and lower **thyroid hormone storage**. This was associated with decreased **embryonic growth**, delayed **hatching** and increased **mortality** during hatching.
OECD TG 206: Avian Reproduction Test	Adult female northern bobwhite quail (*Colinus virginianus*) were exposed to perchlorate at drinking water 1.2, 11.7 and 117 mg/l for 30 d. These doses **did not affect body or organ weights**, or **egg production**, but the top dose caused alterations of **thyroid gland morphology** (NOEC = 11.7 mg/l). These changes included ~50% reduction in **colloid area** and ~30% increase in **follicle cell height**.
Mammalian screens and tests	
TG 408: Repeated Dose 90 Day Oral Toxicity Study	Ammonium perchlorate was administered to rats (n=10) via drinking water at levels of 0.01, 0.05, 0.2, 1.0, and 10.0 mg/kg/day for 90 days. The study design included a non-treatment recovery period of 30 days to evaluate reversibility. The study also investigated potential effects on sperm parameters, estrous cyclicity and serum hormone levels (T3, T4 and TSH). **No toxicologically significant differences** were observed between the control and treated groups with respect to survival, clinical observations, body weights, food consumption, water consumption, haematology, clinical chemistry, **estrous cycling** or **sperm parameters**. In males and females at 10 mg/kg/day **thyroid weights** were increased and **thyroid histopathology** consisting primarily of follicular cell hypertrophy with microfollicle formation and colloid depletion was noted. **These changes were reversible** after a non-treatment recovery period of 30 days. **Changes in TSH and thyroid hormones were observed at all dose levels[1]; however, no thyroid organ weight or histopathological effects were observed at perchlorate dose levels ≤ 1.0 mg/kg/day.**
OECD TG 452-3: Combined Chronic Toxicity/Carcinogenicity Studies	Male Wistar rats were administered sodium perchlorate (1%) via drinking water providing a dose of approximately **1,300 mg/kg/day**. Rats were killed after 40, 120, 220 and 730 days of exposure. Body weights of control and treated animals were similar throughout. **Thyroid weights** of treated rats increased markedly compared to control rats at all time points. After 40 days, the treated rats developed **follicular cell hyperplasia, colloid resorption and low-grade mesenchymal** reaction. After 200 days, **diffuse degenerative changes with fibrosis and increased colloi**d were observed. After 2 years, 4 of 11 treated rats developed **benign thyroid tumours** whilst 20 untreated controls had no thyroid tumours.
	BALBc mice were administered sodium perchlorate (1.2%) in drinking water via drinking water providing a dose of approximately **2,100 mg/kg/day**. **Animals were killed after 46 weeks because of treatment-related deaths**. Perchlorate treatment caused **thyroid epithelial hypertrophy and hyperplasia**. An increased incidence of **thyroid follicular carcinomas** was seen: controls 0/22, perchlorate treated 5/6.

[1] Measurement of thyroid hormone levels is optional in Repeated dose studies (TG 407 and TG 408) and routinely measured in the Extended one-generation reproduction study (TG 443).

6 PROPOSALS BY STAKEHOLDERS AND MEMBER STATES AUTHORITIES

This section provides a comparative overview of the development of proposals for criteria and classification schemes for the assessment of endocrine disrupters under the EU legislative framework that have become public until September 2011. The focus is on general principles and approaches, not on the details of test methods. A compilation of all relevant documents is provided in the final sub-section 6.9 and all references in the text refer to that list.

6.1 ECETOC

Criteria for the identification and assessment of endocrine disrupting chemicals within the framework of European legislation were first proposed by an ECETOC task force of industry scientists. In 2009, the group published an initial guidance document (ECETOC 2009a). In the same year, the proposal was discussed at a workshop with invited scientists from regulatory authorities, industry and academia (ECETOC 2009b, 2009c). As a result, a refined guidance paper was prepared. This became available online in September 2010 and in printed format in 2011 (Bars et al 2011a, 2011b). In May 2011, ECETOC held a second closed workshop on the assessment of endocrine disrupter. The outcome has not yet been published, except for a brief summary on the ECETOC website (Galay-Burgos 2011).

Basically, the refined ECETOC paper suggests six criteria for the assessment of endocrine disrupters in a regulatory context:

- **adversity**,
- **mode of action (MoA)**,
- **causality**,
- **relevance**,
- **specificity**, and
- **potency**.

The first three criteria – adverse effect, endocrine MoA, and evidence for a causal link between both – flow from definitions for endocrine disrupters. The fourth criterion - relevance for humans and/or wildlife populations – reflects the limitations of evidence coming from laboratory test methods. The other two criteria - considerations of specificity and potency - are suggested to come on top.

The paper discusses the range of currently available test methods that may provide relevant *in vitro*, *in viv*o screening and *in vivo* apical data. With many separate assays available, the authors stress the need for a holistic evaluation of all available data. To this end, "*an objective, systematic and structured* ***weight-of-evidence evaluation***" is considered to be highly important to establish conclusive proof of endocrine disruption.

If there is sufficient evidence for a substance to be considered an endocrine disrupter, the authors propose a consideration of the **specificity** of endocrine effects as a subsequent assessment step. In the context of human health protection, the authors argue that an "*assessment of specificity is*

required to determine whether the adverse effects observed occur at dose levels lower than other forms of toxicity, e.g. neuro-, hepato- or cardio-toxicity. If this is not the case then risk assessment of the substance should be based on the most sensitive non-endocrine endpoint". In the context of assessments for wildlife protection, additionally the aspect of species specificity comes into play. The authors explain "*that consideration may also be given to the specificity of endocrine effects in relation to general toxic endpoints in other taxonomic groups which may drive the overall risk assessment. For example, an endocrine effect in fish may occur at concentrations above those inducing general toxic effects in other species (e.g. algae) affording it lower concern considering the inbuilt margin of safety incorporated into the risk assessment*".

If a substance has been demonstrated to be an endocrine disrupter, if the effects are specific and relevant to humans and/or wildlife populations, and if the exposure is not negligible according to the rules of a specific piece of legislation, such as the plant protection product regulations, then the authors propose **potency** considerations as a final assessment step. Several descriptors of potency should be collectively taken into consideration: "*the dose or concentration at which adverse effects are caused, the duration of exposure that is required for an adverse effect to be induced, the type, incidence and severity of the effect, as well as the number of species in which adverse effects were demonstrated.*"

The aim of the potency assessment is to discriminate between endocrine disrupters of high concern and those of lower concern. However, in contrast to the proposals brought forward by competent German and UK authorities (*see below*), the ECETOC work does not suggest any specific potency-based cut-off values for regulatory decision making. The authors consider hazard to be inappropriate as a sole criterion and argue for risk assessments of endocrine disrupters with assessment factors based on potency.

6.2 UK CRD

The UK Chemicals Regulation Directorate (CRD) has developed two draft papers for the regulatory definition of endocrine disrupters. They are based on UK government-wide consultations and endorsements, including the Committee on Toxicity (COT 2010). The first paper deals with the definition in relation to potential threat to human health and was communicated in December 2010 (CRD 2010). The counterpart paper sets out criteria for ecotoxicological endocrine disrupters and became available in April 2011 (CRD 2011b). In May 2011, the paper related to human health became superseded by a joint German-UK proposal (see section 6.4), while the ecotoxicological counterpart paper still reflects the position of the agency.

Human health

The human health related paper considers that the identification of a substance as an endocrine disrupter "*is potentially of great regulatory and commercial impact*". "*Hence this paper takes the position that the assigning of the ED identifier to a substance should be reserved for those substances where such a property is clearly established, the substance is potent in this respect, and the endocrine-disrupting property is a dominant feature of the hazard profile of the substance*" (paragraph 6). The WHO/IPCS definition is considered to be a very broad description that has no power for discriminating between endocrine disrupters of high concern and those of low concern. As

a consequence, the paper aims "*to use the WHO definition as the starting point to arrive at a regulatory definition of an ED by adding a number of criteria that need to be satisfied before an ED requiring regulatory action can be identified*" (paragraph 17).

The paper argues for the use of the same set of criteria as in the ECETOC proposal (*see above*), with the only difference that a slightly altered wording is used for what the ECETOC paper refers to as "*specificity*". The CRD proposal calls this criterion "*most sensitive/**lead toxic effect**(s)*" or "*dominant feature of the hazard profile".* However, an important difference comes from the fact that the CRD paper transforms the potency criterion into definite **potency-based cut-off values** that mark the borderline between substances that are regarded as EDs for regulatory purposes and those which are not. To this end, the paper proposes to make use of dose thresholds for STOT-RE classification (Specific Target Organ Toxicity – Repeated Exposure) which have been defined for the purpose of classification, labelling and packaging of hazardous substances in the corresponding European regulation and the Globally Harmonised System (GHS). It "*is suggested that only where a substance produces endocrine disruption at a dose level at or below the discriminatory guidance dose levels for the application of Category 1 STOT-RE hazard classification, should the substance remain under consideration as a potential ED for regulatory purposes*" (paragraph 32). In the case of data from a 90 days study with oral exposure for instance, this cut-off value would be equivalent to a dose level of 10 mg/kg bw/day.

In summary, the paper proposes that a substance is regarded as an ED for regulatory purposes if it satisfies the following criteria (paragraph 43):

- *"adverse effects to 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.*
- *the adverse effect(s) believed to be related to endocrine disruption to be the lead toxic effect(s) in the dataset; or occurring at a dose level close to that at which the lead toxic effect was first seen.*
- *serious adverse effect(s) believed to be related to endocrine disruption to 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.*
- *a mode-of-action link between the toxic effects of concern and endocrine disruption to have been established.*
- *the effects seen in experimental animals to be judged to be of potential relevance to human health."*

As mentioned above, this draft position paper was superseded by a joint German-UK proposal in May 2011. The major difference between both versions is that the concept of the *lead toxic effect(s)* was dropped (*see section 6.4*).

Wildlife

The counterpart paper related to wildlife follows the same line of argumentation and criteria as the version for human health effects with the following differences and specifications:

- **No potency-based cut-off values** are suggested.

- The criterion of the ***lead toxic effect(s)*** is applied analogously, but the terms *lead toxic effect(s)* or *specificity* (ECETOC) are not used. Instead the criterion is described by means of an example:
  "*In determining whether a compound is an ecotoxicological endocrine disruptor or not there should be a consideration of the concentration or dose causing ED effects. For example, if the key endpoint from fish assays and the full fish life cycle study are several orders of magnitude greater than other key endpoints then the ED effect can be considered to be of limited regulatory relevance. This is due to the fact that any regulatory decision, for example no authorisation, implementation of buffer zones or other risk mitigation measures will be based on a significantly lower endpoint. This is illustrated by the following example – a new herbicide has an EC50 for Lemna of 1.0 µg a.s./L. This is the lowest endpoint and is 'driving' the risk assessment whereas the NOEC from the full fish life cycle study is 10 mg a.s./L. In this situation, it is proposed that the results of the FFLC are of limited regulatory relevance*" (paragraph 33).
- The focus is on the protection of populations. Therefore a distinction is made between findings in laboratory studies with **population relevance** and effects that are not likely to affect the population recruitment and stability. This is in line with the ECETOC proposal.
- **Only guideline studies** are considered to provide relevant information for regulatory identification of an ecotoxicological endocrine disrupter: "*All studies used in an assessment must have been conducted to an internationally recognised protocol*" (paragraph 24)
- The paper points out that the assessment of endocrine properties of a substance needs to be pursued separately for each major category of animals in the environment. However, for several important groups of species internationally agreed **test guidelines are missing**. Given these constraints, the paper concludes that "*it is really only possible to determine whether or not a substance is an ecotoxicological ED for mammals and fish*" (paragraph 29). For birds, amphibians, reptiles, and invertebrates the available suite of assays is judged to be inadequate for achieving conclusive evidence.
- Some **substances designed to act as an endocrine disrupter** in certain target invertebrate species are used as active ingredients of plant protection and biocidal products, e.g. so-called insect growth regulators. For these special compounds, the paper proposes that "*investigations should be undertaken to explore whether or not there is an adverse effect at the population level and at the field scale. Where such findings arise, then it might be appropriate to conclude that a substance is an ED in relation to non-target invertebrates in the environment*" (paragraph 43).

In summary, the paper concludes that a substance should be regarded as an ED for regulatory purposes if it satisfies the WHO/IPCS definition and the following criteria (paragraph 44):

- **"***the nature of the effect must pose a threat to population recruitment or stability: and*
- *there should be a reasonable and coherent line of evidence from within the same taxonomic group that the mode-of-action underlying the effect observed is endocrine disruption*
- *there should be a consideration of the concentration/dose causing adverse endocrine effects as the example described in paragraph 33*" (*see above*).

6.3 GERMAN BFR, BAUA AND UBA

Concerning the protection of human health, separate proposals for the assessment of endocrine disrupters under the plant protection product regulation (PPPR) and for the identification of endocrine disrupters as substances of very high concern (SVHC) under REACH were initially prepared by the German Federal Institute for Risk Assessment (BfR 2010a) and the German Federal Institute for Occupational Safety and Health (BAuA 2010), respectively. The BfR proposal was based on discussions at an expert workshop held at the BfR in November 2009 (BfR 2010b) and it was later also published in the open scientific literature (Marx-Stoelting et al. 2011). Both the initial BfR proposal and the BAuA paper became superseded with the publication of a joint DE – UK position paper in May 2011 (BfR 2011) which covers human health related assessments under both the PPPR and REACH, and additionally under the regulation for biocidal products (*see section 6.4*).

Concerning the protection of wildlife, the competent German Federal Environment Agency (UBA) did not yet present a fully worked out proposal, comparable in scope and details to the counterparts in the human health arena (BAuA 2010, BfR 2010, 2011) or the British proposal for the regulatory definition of an ecotoxicological endocrine disrupter (CRD 2011b). However, contributions to the discussion about assessment criteria for endocrine disrupters both under the PPPR (UBA 2010a) and under REACH (UBA 2010b) were communicated to EU Member States.

Human health

The initial proposals by both the BfR (2010) and the BAuA (2010) suggested the same criteria as the corresponding UK CRD proposal (CRD 2010). In particular, the same STOT-RE Cat 1 values were suggested to be used as a **potency-based cut-off** criterion under the PPR and as a trigger value for the SVHC identification under REACH. The BAuA paper defined this as a pragmatic "**threshold of regulation**" (page 6).

However, there is one important difference between these initial German proposals and the British counterpart: The **German agencies did not adopt the criterion of the "lead toxic effect"** as it was named in the CRD proposal and which was originally proposed under the term "specificity" by ECETOC. The BAuA paper explicitly stated that for SVHC identification it "*is not a pre-requisite that the ED-related adverse health effect is the most sensitive adverse effect that has been identified and/or that was the lead effect driving classification (e.g. the same substance might also be neurotoxic at a much lower dose, which is not (yet) specified as a criterion for SVHC identification)*" (page 7-8). Consequently, this criterion was later also totally abandoned in the joint DE-UK position paper (BfR 2011).

In the initial BfR discussion paper (BfR 2010), the potency-based classification by means of the STOT-RE Cat 1 categorisation values was considered as one of two possible options for ED assessments under the PPPR. It was named the "*classification-based option*". Alternatively, an "*exposure-based option*" was proposed. This meant that an exposure analysis should "*be performed to find out whether or not exposure to the respective substance for consumers as well as for operators, workers and others who might be exposed to the substance is negligible*" (page 10). To this end the following definition of "*negligible exposure*" of workers, consumers and bystanders to an active substance was proposed and explained in Annex I to the BfR paper:

a.) "*total systemic exposure or local exposure counts for less than 10% of the corresponding reference value (AEL) or*
b.) *the active substance is not genotoxic and the total internal exposure to the active substance does not exceed 1.5 µg per person and day*".

However, with the publication of the joint DE-UK paper this option is obviously no longer pursued.

Wildlife protection under the PPPR

The UBA paper on the assessment of substances with endocrine disrupting properties within the framework of environmental risk assessments conducted under the PPPR was communicated in May 2010 (UBA 2010a). The paper argues for differentiated decision-making in the following three types of assessment situations:

A. *"The endocrine disrupting properties are the basis of the pesticidal mechanism of action in target organisms (only invertebrates and plants)*".
B. "*The endocrine disrupting properties are decisive for the overall side-effects on non-target organisms*".
C. "*The endocrine disrupting properties are not decisive for the overall side-effects on non-target organisms*".

The following three types of information were considered as a basis for decision making:

i. *"Specificity of the endocrine disrupting properties"*
ii. *"Potency of endocrine disrupting properties"*
iii. *"Ecological relevance/adversity of endocrine mediated effects"*

For situation A, i.e. substances that were designed to act as endocrine disrupters in target organisms, it is proposed to perform established risk assessment procedures by way of non-compliance with the hazard-based cut-off criterion established under the PPPR. This differs from the corresponding CRD proposal which suggested conducting field studies in this situation (see section 6.2).

The differentiation between situations B and C, i.e. "*decisive*" or "*not decisive for the overall side-effects*" is equivalent to the criterion "*specificity*" (ECETOC) or "*lead toxic effect*" (CRD). It is proposed that the hazard-based cut-off criterion of the PPPR should only apply to situation B ("*decisive for the overall side-effects*"), otherwise (situation C) substances should undergo established risk assessment procedures. This is in agreement with the CRD proposal for ecotoxicological assessments; it is different from the assessment schemes proposed by BAuA and BfR for hazards to human health, where the concept of the "*lead toxic effect*" was rejected (*see above*).

Wildlife protection under REACH

The UBA discussion paper on the assessment of endocrine disrupters under REACH was also communicated in May 2010 (UBA 2010b). The paper describes general thoughts on how to identify substances that give rise to an equivalent level of concern to those of CMRs, PBTs or vPvBs, due to their endocrine disrupting properties for organisms in the environment. The purpose was to initiate

a discussion among Member States, ECHA, European Commission and the Member State Committee (MSC).

The paper considers the WHO/IPCS definition of endocrine disrupters and concludes that "*a substance should fulfill at least two criteria in order to be identified as an endocrine disruptor for organisms in the environment:*

- *endocrine mode of action*
- *adverse effects in the environment*" (page 3).

The authors point to the difficulties that may arise in demonstrating a causal link between an endocrine MoA and an observed adverse effect and therefore they consider "*all available information*" to be "*relevant in order to identify endocrine disruptors in a weight of evidence approach*" (page 3). This view differs from the CRD paper on ecotoxicological endocrine disrupters which argues that regulatory decisions should be exclusively based on so-called guideline studies (see section 6.2).

The paper further considers possible interpretations of REACH Art. 57 (f) under guidance of the precautionary principle as well as the potentially available types of data. As a result, the authors come to conclude that "*a substance should be identified as of equivalent very high concern for the environment*" if it satisfies the definition (see above) and in doing so has:

- *"a high impact to the environment (i.e. population health), although risk cannot be determined with sufficient certainty*,
- *the potential to cause irreversible impact on further generations*" (page 6).

6.4 JOINT DE – UK POSITION PAPER

Following detailed comments from the UK CRD on the German BAuA paper (CRD 2011a), the German BfR and the UK CRD worked out a joint position paper on the regulatory definition of an endocrine disrupter in relation to potential threats to human health. It was published in May 2011 (BfR 2011). This document supersedes the corresponding previous proposals of BfR (2010), BAuA (2010) (*see section 6.3*) and the CRD (CRD 2010) (*see section 6.2*).

The text of the joint paper is an amended version of the previous CRD paper (CRD 2010). The important change is that the requirement for the adverse endocrine-mediated effect to be also the **"lead toxic effect" has been dropped**. In addition, the requirement for the demonstration of a mechanistic link between an endocrine MoA and an adverse effect in an intact organism was reformulated in a slightly less restrictive way. The requirement for "*a mode-of-action link*" was rephrased to "*a plausible mode-of-action/mechanistic link*". Apart from these amendments and a change in the sequence of arguments, the text of the concluding summary of criteria remained unchanged. Accordingly, it is proposed that a substance is regarded as an "*ED of very high regulatory concern*" when it satisfies the WHO/IPCS definition and in doing so satisfies the following criteria (paragraph 40):

- "*adverse effects to 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.*

- *a plausible mode-of-action/mechanistic link between the toxic effects of concern and endocrine disruption.*
- *the effects seen in experimental animals to be judged to be of potential relevance to human health.*
- *serious adverse effect(s) related to endocrine disruption to 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 & labeling*".

6.5 FRENCH ANSES

The French Agence Nationale de Sécurité Sanitaire de l´alimentation, de l´environnement et du travail (ANSES) has not elaborated a detailed separate proposal for ED identification and categorisation. However, ANSES communicated written comments to the German BAuA's proposal regarding human health criteria for endocrine disruption (see section 6.3).

The main point of ANSES' critique is the introduction of a potency-based cut-off by means of the STOT-RE Cat 1 trigger values. The authors explain their strong reservations against this strategy.

In summary, ANSES considers the following two parameters to be sufficient for identifying an endocrine disrupter as an SVHC according to REACH article 57(f):

- *"Classification for Carcinogenicity or Reproductive toxicity (cat 1A, 1B or 2) or STOT-RE (cat 1 and cat 2), based on human and in vivo data through physiological routes of exposures as proposed in the classification criteria and guidance.*
  and
- *Identification of an ED mode of action for the adverse effect leading to the classification(s) mentioned above (based on a weight of evidence also using screening tests and in vitro tests)*" (page 3).
- The inclusion of both STOT-RE cat 1 and cat 2 means that the potency-based trigger value is considerably raised in comparison to the joint DE – UK position paper which suggests considering only a STOT-RE cat 1 substance as an "*ED of very high regulatory concern"*. In the case of data from a 90 days study with oral exposure for instance, the resulting cut-off value would be increased from 10 to 100 mg/kg bw/day.

6.6 DANISH EPA

The Danish Environmental Protection Agency has (i) given detailed written comments to the proposals of other EU Member States in the years 2010 and 2011 (DK EPA 2010a, 2010b, 2011c), (ii) prepared a detailed discussion paper on the regulation of EDs under REACH (DK EPA 2011b), (iii) commissioned a scientific report on criteria for endocrine disrupters (Hass et al. 2011), and (iv) finally published a comprehensive proposal for the "*Establishment of criteria for endocrine disruptors and options for regulation*" in May 2011 (DK EPA 2011a).

The three most important differences between the Danish proposal and the other proposals outlined in the preceding sections are:

- no potency-based cut-off criteria,
- no requirement for the endocrine mediated adverse effect being the "*lead toxic effect*", neither in humans nor in the context of ecotoxicological assessments,
- slightly less restrictive and more precautious requirements regarding the evidence that is considered to be sufficient for concluding that a substance satisfies the WHO/IPCS definition of endocrine disrupters, in particular with respect to the proof of a causal link between an endocrine MoA and an observed adverse effect with relevance in humans or wildlife populations.

In compliance with the WHO/IPCS definition, the Danish proposal basically subdivides endocrine disrupters in two categories: "*confirmed*" EDs (Cat 1) (based on *in vivo* data*)* and "*potential*" EDs (Cat 2). However, for regulatory purposes it is proposed to further subdivide the "*potential*" EDs in Cat 2 into two subgroups reflecting the level of available evidence: "*suspected*" EDs (Cat 2a) (mainly based on *in vivo* data) ; and "*indicated*" EDs (Cat 2b), i.e. "*substances with indications of ED properties*" (mainly based on *in vitro/in silico* data*)*.

For Cat 1 EDs no approval shall be granted under the PPPR unless exposure is negligible as defined in the regulation, and Cat 1 EDs shall be treated as SVHCs under REACH. For Cat 2 EDs other regulatory actions may apply as outlined in the following scheme.

Schematic outline of the Danish proposal for definitions and regulatory options (from DK EPA 2011a, page 11)

WHO/IPCS definition	***EU definition***	***REACH – possible regulatory actions***	***PPPR – regulatory action***
ED	*Category 1: ED (in vivo data)*	• *Identification as SVHC (and possible inclusion in the authorisation list, Annex XIV)* • *Restriction* • *Harmonised C&L*	*No approval unless negligible exposure*
Potential ED	*Category 2a: Suspected ED (mainly in vivo data)*	• *Development of list of potential EDs* • *For prioritised potential EDs, development of RMO analysis followed by regulation, if appropriate* • *Prioritisation for substance evaluation where more data on ED specific properties can be required from industry* *	*Approval requires further data from industry*
	Category 2b: Indicated ED (mainly in vitro/in silico data)	**	*Depending on the case, flag for generation of further data*

** For suspected EDs (category 2a): Generation of further ED specific data can be conducted by industry, Member States and research communities on a voluntary basis.*

*** For indicated EDs (category 2b): Generation of further data to be prioritised depending on exposure*

The Danish proposal provides detailed criteria for every suggested category. For Cat 1 substances these are the following (DK EPA 2011a, page 6, Table 1):

"*Substances are placed in category 1 when they are known to have caused ED mediated adverse effects in humans or animal species living in the environment or when there is evidence from animal studies, possibly supplemented with other information, to provide a strong presumption that the substance has the capacity to cause adverse ED effects in humans or animals living in the environment.*

The animal studies shall provide clear evidence of ED effects in the absence of other toxic effects, or if occurring together with other toxic effects, the ED effects should be considered not to be a secondary non-specific consequence of other toxic effects. However, when there is e.g. mechanistic information that raises doubt about the relevance of the effect for humans or the environment, category 2a may be more appropriate.

Substances can be allocated to this category based on:

- *Adverse in vivo effects where an ED mode of action is highly plausible*
- *ED mode of action in vivo that is clearly linked to adverse effects in vivo (by e.g. read-across)*".

Additionally, the evidence required for fulfillment of criteria is defined in terms of data resulting from assays compiled in the revised OECD Conceptual Framework (OECD CF) for endocrine testing and assessment. Accordingly, a "*substance can be considered a confirmed ED (category 1) based on data from:*

- *In vivo assays providing data on adverse effects clearly linked to endocrine mechanisms (OECD CF3, level 5)*
- *On a case-by-case basis, in vivo assays providing data about single or multiple endocrine mechanisms and adverse effects (OECD CF, level 3 & 4) combined with other relevant information*
- *In special cases, where in vivo data on adverse effects are lacking, categorisation or (Q)SAR approaches may provide the necessary data in combination with ADME information and in vitro data*
- *Reliable and high quality evidence from human cases or epidemiological studies*" (page 7).

6.7 CHEM TRUST

The environmental NGO CHEM Trust, in co-operation with WWF, has (i) published a detailed critique of the proposals and position papers of the British, German, Danish, and French governmental agencies (CHEM Trust 2011), (ii) led the preparation of a joint position paper of twelve environment and health NGOs on "*Requirements for the proper regulation of chemicals with endocrine disrupting properties*" (CHEM Trust et al. 2011), and (iii) published a discussion paper that suggests a possible classification scheme for chemicals with ED properties (CHEM Trust & WWF 2010).

CHEM Trust and WWF propose that a practical scheme for tackling EDC chemicals would be to introduce four subcategories (1A, 1B, 1C or 2) as outlined in the following scheme (from CHEM Trust & WWF 2010, page 16):

Summary of proposed categorisation scheme for EDCs (E)	***E2***	***E1C***	***E1B***	***E1A***
Substances suspected of being EDCs on the basis of simple in-vitro tests or non-validated QSARs (which don't take account of metabolism) – unless other data negate concerns.	*x*			
Substance considered to have ED properties in vivo e.g. cause effects on hormone levels, or hormone sensitive tissues, or endocrine glands, or auxiliary systems.		*x*	*X*	*X*
Some suspicion of endocrine mediated effects – or endocrine disruption known/strongly suspected but unsure if effects are adverse (e.g.: substance tests positive in the uterotrophic or Hershberger test.)		*x*	*X*	*X*
Strong suspicion of endocrine mediated effects.			*X*	*X*
Evidence to show effect is direct consequence of disruption of endocrine system. Causal mechanism known with certainty.				*X*

The paper provides a detailed proposal on how to use the categorization scheme under REACH and the PPR respectively. In general, it is suggested that Cat 1 EDs (including 1A, 1B and 1C) should be subject to authorization under REACH and should be phased out of use under the PPPR unless they give rise to negligible exposure. Exemptions may apply to some pesticides that have been specifically designed to achieve selective target organism toxicity via endocrine modes of action. This is in line with the corresponding proposal of the German UBA (see section 6.3).

6.8 PAN EUROPE

The European branch of the international Pesticide Action Network (PAN Europe) participated in the joint position paper of twelve environment and health NGOs that was prepared under the leadership of CHEM Trust in April 2011 (CHEM Trust et al. 2011). In May 2011, PAN Europe raised harsh criticism of the joint DE - UK position paper (BfR 2011) in an open letter to the European Commissioners for Health and Consumer Policy and for the Environment (PAN Europe 2011a). This was accompanied by the publication of a PAN Europe position paper on *criteria for the determination of endocrine disrupting properties of pesticides* (PAN Europe 2011b).

In contrast to proposals from EU Member States authorities, the PAN Europe position paper does not accept the scientific WHO/IPCS definitions as a basis for regulatory consensus finding, neither the definition of endocrine disrupters nor the definition of adverse effects. Furthermore, it rejects all additional criteria under discussion, not only the concept of the *lead toxic effect* and *potency* based cut-off criteria, but also considerations of *relevance* for humans or wildlife populations. And in contrast to all other proposals, it does not consider any sub-categorisations such as *confirmed* and *potential* EDs for instance.

The PAN Europe paper puts strong emphasis on the exact wording of the cut-off criterion laid down in Regulation (EC) 1107/2009 on plant protection products which says that active substances, safeners and synergists shall only be approved, if they are not considered to have "*endocrine disrupting properties that may cause adverse effects"* in humans or non-target organisms. As a consequence, the authors consider discussions about a definition of an "*endocrine disrupter*" or an "*endocrine disrupting chemical*" to be "*futile*" and they state that "*mentioning of definitions like the one of WHO/IPCS is unwanted and ... (…) changing democratically agreed policy"* (p. 3).

For classifying a chemical as having "*endocrine disrupting properties"*, the paper considers "*any effect on the endocrine system"* to be a sufficient criterion, including both direct and indirect effects, and the paper states that for this classification "*known mechanisms of action should not be necessary*" (p. 6). For the complementary definition of "*adverse effect*", the paper proposes "*to take any significant biochemical alterations, following dosing with the chemical during key development stages that is above background or averages in testing*" (p. 7).

Concerning the required strength of evidence, the paper states that "*it should be possible to classify a chemical as having endocrine disrupting properties on the basis of a positive in-vitro study (…), unless within a specific set time-period industry has come up with in-vivo data which shows without any doubt the in vitro results to have been a false positive*" (p. 6).

Regarding data requirements, the paper calls for the development of a test battery that includes "*methods for ALL hormonal systems"* and a special test system that insures exposure of test animals

during special windows of vulnerability, that addresses all potential endocrine disruption related endpoints, and that includes tests at low doses for detecting non-monotonic dose response curves. The paper argues that "*independent scientists … (….) should be given a leading position in developing tests*" (p. 4) and it suggests that the "*Joint Research Centre (JRC) of the EC could take the lead in establishing a forum for developing tests"* (p. 5).

6.9 SUMMARY OF MAJOR CONTROVERSIAL ISSUES

All proposals, except the PAN Europe position paper, more or less accept the WHO/IPCS definition of endocrine disrupters as a suitable working basis for the development of regulatory criteria for the identification and assessment of endocrine disrupters. Beyond this definitional issue, the comparative overview shows that three major controversial points dominate the current debate:

- the use of potency-based STOT-RE Cat 1 trigger values as cut-off criteria for endocrine disrupters of regulatory concern, which is favoured by the joint DE-UK proposal but categorically rejected by France and Denmark,
- the concept of the "*lead toxic effect*" which has been abandoned in member state proposals relating to the assessment of human health effects but which remains under discussion for the purpose of ecotoxicological assessments,
- the strength of evidence that should be regarded as sufficient to assume that a substance satisfies the WHO/IPCS definition for regulatory purposes.

6.10 DOCUMENT LIST

ANSES (Agence nationale de sécurité sanitaire de l´alimentation, de l´environnement et du travail) (France) (2010) Comments further to BAuA document on human health criteria for endocrine disruption (ED) according to Art. 57 (f) of the REACH regulation. Maisons-Alfort, dated 10/01/2011 (unpublished communication)

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, Galay-Burgos M (2011a) Science based guidance for the assessment of endocrine disrupting properties of chemicals. Regul Toxicol Pharmacol 59, 37–46

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, Galay-Burgos M (2011b) Corrigendum to "Science based guidance for the assessment of endocrine disrupting properties of chemicals" [Regul Toxicol Pharmacol 59 (2011), 37–46]. Regul Toxicol Pharmacol 60 (2011), 411-12

BAuA (Bundesanstalt für Arbeitsschutz und Arbeitsmedizin) (German Federal Institute for Occupational Safety and Health) (2010) Human health criteria for endocrine disruption (ED) according to Art. 57 (f) of the REACH Regulation: German approach to the identification of ED substances as SVHC. Dortmund, dated 20 October 2010. Confidential, to DG ENV, ECHA and MSCA. Available in the public domain via: http://chemicalwatch.com/downloads/DE_paper_-_ED_criteria_HH_Caracal_6%5B1%5D.pdf

http://chemicalwatch.com/downloads/DE_paper_-_ED_criteria_HH_Caracal_6%5B1%5D.pdf

BfR (Bundesinstitut für Risikobewertung) (German Federal Institute for Risk Assessment) (2010a) Draft concept paper. Development of a stepwise procedure for the assessment of substances with endocrine disrupting properties with according to the plant protection products regulation (Reg. (EC) No 1107/2009). Dated 05 Mai 2010. http://www.bfr.bund.de/cm/343/development_of_a_stepwise_procedure_for_the_assessment_of_substances_with_endocrine_disrupting_properties.pdf

BfR (Bundesinstitut für Risikobewertung) (German Federal Institute for Risk Assessment) (2010b) Establishment of assessment and decision criteria in human health risk assessment for substances with endocrine disrupting properties under the EU plant protection product regulation. Report of a Workshop hosted at the German Federal Institute for Risk Assessment (BfR) in Berlin, Germany, from Nov. 11th till Nov. 13th 2009. Federal Institute for Risk Assessement, Berlin, Germany, January 15th, 2010. http://www.bfr.bund.de/cm/343/establishment_of_assessment_and_decision_criteria_in_human_health_risk_assessment_for_substances_with_endocrine_disrupting_properties.pdf

BfR (Bundesinstitut für Risikobewertung) (German Federal Institute for Risk Assessment) (2011) Joint DE – UK position paper. Regulatory definition of an endocrine disrupter in relation to potential threat to human health. Proposal applicable in the regulatory context of Plant Protection Products, Biocidial Products, and Chemicals targeted within REACH. Dated 16th Mai 2011. http://www.bfr.bund.de/cm/343/regulatory_definition_of_an_endocrine_disrupter_in_relation_to_potential_threat_to_human_health.pdf

CHEM Trust (2011) CHEM Trust's contribution to the ongoing debate on criteria for EDCs. Paper developed with input from the WWF European Policy Office. Dated September 2011. http://www.chemtrust.org.uk/documents/CHEM%20Trust%20Position%20on%20EDC%20Criteria%20-%20Sept11.pdf

CHEM Trust, Cancer Prevention & Education Society, WECF, Health Care Without Harm, Pesticide Action Network Europe, EEB, ChemSec, Health & Environment, Alliance, Global 2000, Bund, BEUC and istas (2011) Environment and health NGOs', consumer organisations & trade union's position paper, April 2011. Requirements for the proper regulation of chemicals with endocrine disrupting properties. http://www.chemtrust.org.uk/documents/NGOs%20requirements%20EDCs%20April%202011-FINAL.pdf

CHEM Trust, WWF (2010) Protecting future generations by reducing exposure to endocrine disruptors. CHEM Trust and WWF-EPO proposals for the regulation of chemicals with endocrine disrupting properties under REACH (EC 1907/2006) and under the Plant Protection Products Regulation (EC No 1107/2009). A joint discussion paper by CHEM Trust and WWF European Policy Office dated December 2010 (updating an initial WWF discussion paper of 2002) http://www.chemtrust.org.uk/documents/CHEM%20Trust%20&%20WWF%20EDC%20Classification%20Paper%20Dec%202010.pdf

COT (Committee on toxicity of chemicals in food, consumer products and the environment) (2010) Regulatory definition of endocrine disrupter. This is a draft paper for discussion, it does not reflect views of the Committee. It should not be quoted, cited or reproduced. COT TOX/2010/15. Available online at http://cot.food.gov.uk/pdfs/tox201015.pdf

CRD (UK Chemicals Regulation Directorate) (2010) Regulatory definition of an endocrine disrupter in relation to potential threat to human health. Draft, electronic file dated Dec 2010 (unpublished communication)

CRD (UK Chemicals Regulation Directorate) (2011a) Comments from the UK REACH CA on the German paper on the identification of an ED for the purpose of Art 57(f). (undated unpublished communication; communicated to CARACAL 03 Dec 2010)

CRD (UK Chemicals Regulation Directorate) (2011b) Definition of an ecotoxicological endocrine disrupter for regulatory purposes. Electronic file dated Apr '11 (unpublished communication)

DK EPA (Danish Ministry of the Environment, Environmental Protection Agency) (2010a) Comments to the BAuA document: "Human health criteria for endocrine disruption (ED) according to Art. 57 (f) of the REACH regulation: German approach to the identification of ED substances as SVHC", 20th October 2010. Dated 30th November 2010. Published as Chapter 4 of Annex C to Danish EPA 2011a

DK EPA (Danish Ministry of the Environment, Environmental Protection Agency) (2010b) Comments to the UBA document: "Discussion paper on interpretation of Art. 57(f) REACH with respect to substances having endocrine disrupting properties hazardous to the environment", May 26th, 2010. Dated 12th of July 2010. Published as Chapter 5 of Annex C to Danish EPA 2011a

DK EPA (Danish Ministry of the Environment, Environmental Protection Agency) (2011a) Establishment of criteria for endocrine disruptors and options for regulation. Dated 17 May 2011.
http://www.mst.dk/NR/rdonlyres/54DB4583-B01D-45D6-AA99-28ED75A5C0E4/127098/DKEDcriteria110517_finalcorr1.pdf

DK EPA (Danish Ministry of the Environment, Environmental Protection Agency) (2011b) Regulation of endocrine disruptors under REACH. Dated 31 January 2011. Published as Annex B to Danish EPA 2011a

DK EPA (Danish Ministry of the Environment, Environmental Protection Agency) (2011c) Danish comments to input from other regulatory bodies/member states in relation to assessment of endocrine disrupters. Dated 17 May 2011. Published as Annex C to Danish EPA 2011a

ECETOC (European Centre for Ecotoxicology and Toxicology of Chemicals) 2009a. Guidance on identifying endocrine disrupting effects. ECETOC Technical Report No. 106, Brussels, Belgium, ISSN-0773-8072-106

ECETOC (European Centre for Ecotoxicology and Toxicology of Chemicals) (2009b). Workshop: Guidance on identifying endocrine disrupting effects, 29–30 June 2009, Barcelona. ECETOC Workshop Report No. 16, Brussels, Belgium, October 2009.

ECETOC (European Centre for Ecotoxicology and Toxicology of Chemicals) (2009c). Addendum to ECETOC Workshop Report No. 16. Brussels, Belgium, 22 December 2009. Notice provided online at www.ecetoc.org

Galay-Burgos M (2011) Risk assessment of endocrine disrupting chemicals (ECETOC workshop), 9-10 May 2011, Firenze, Italy (brief information on workshop structure and outcome), provided online at
HTTP://WWW.ECETOC.ORG/INDEX.PHP?MACT=XTENDEDCALENDAR,CNTNT01,DETAILS,0&CNTNT01DOCUMENTID=38&CNTNT01RETURNID=108

Hass U, Christiansen S, Boberg J, Vingaard AM, Andersson A-M, Skakkebæk NE, Bay K, Holbech H, Bjerregard P (2011) Report on criteria for endocrine disrupters. Danish Centre on Endocrine Disrupters, dated May 2011. Published as Annex A to Danish EPA 2011a

Marx-Stoelting P, Pfeil R, Solecki R, Ulbrich B, Grote K, Ritz V, Banasiak U, Heinrich-Hirsch B, Moeller T, Chahoud I, Hirsch-Ernst KI (2011) Assessment strategies and decision criteria for pesticides with endocrine disrupting properties relevant to humans. Reprod Toxicol 31, 574-84

PAN Europe (Pesticide Action Network Europe) (2011a) Letter to Commissioner Janez POTOCNIK, concerning criteria for endocrine disrupting pesticides. Brussels, 25-05-2011.
http://www.pan-europe.info/News/PR/110525_Letter%20Mr.%20Potocnik%20criteria%20for%20ED%20pesticides.doc

PAN Europe (Pesticide Action Network Europe) (2011b) Criteria for the determination of endocrine disrupting properties. PAN Europe position paper, Brussels, 25-05-2011. http://www.pan-europe.info/News/PR/110525_Criteria%20for%20ED%20PAN%20Europe%20position.doc

UBA (Umweltbundesamt) (German Federal Environment Agency) (2010a) Substances with endocrine disrupting properties in the environmental risk assessment under the new EU regulation on plant protection products (EC1107/2009) – A Proposal for a differentiated decision making. Dr Tobias Frische, dated 2010-01-05 (unpublished communication)

UBA (Umweltbundesamt) (German Federal Environment Agency) (2010b) Discussion paper on the interpretation of Art 57(f) REACH with respect to substances having endocrine disrupting properties hazardous to the environment. Frauke Stock, Dessau-Roßlau, dated May 26th 2010 (unpublished communication)

7 TESTING REQUIREMENTS AND DECISION CRITERIA: OPTIONS AND RECOMMENDATIONS

It is widely acknowledged that current testing practice in the context of CLP, REACH or PPPR is not designed for the identification of endocrine disrupters, although certain endpoints and assays may give some indication of endocrine disrupting properties (see section 5). In general though, testing for endocrine disrupting (ED) properties with the most appropriate endpoints and with exposure regimes that cover periods of susceptibility during development is not yet conducted. Under REACH, it is also not mandatory.

When dealing with endocrine disrupters in the regulatory context, two related issues must be addressed:

First, appropriate tests and endpoints for the identification of ED properties need to be implemented in the context of relevant EU pieces of legislation, namely REACH and PPPR.

Secondly, criteria for the translation of test outcomes into regulatory decisions need to be developed.

7.1 IMPLEMENTATION OF TESTS FOR THE IDENTIFICATION OF CHEMICALS WITH ENDOCRINE DISRUPTING PROPERTIES - REQUIREMENTS AND TESTING STRATEGIES

This section considers general principles and issues important for the implementation of tests and assays that are required for the identification of substances with ED properties. However, it is far beyond the scope of this report to make detailed proposals in terms of the precise assays to be chosen, the sequence in which testing should be conducted, and the waiving of testing needs. That will have to be the topic of further detailed studies and discussions among experts in competent authorities, and will have to include the development of detailed guidance documents for the interpretation of data.

In view of their different regulatory aims, protection goals and regulatory set up, PPPR and REACH are considered separately.

In general, however, the implementation of tests will have to rely on validated and internationally agreed test methods. It should be recognised, that this **limits very considerably the range of endocrine disrupting effects on which regulatory decisions can currently be based**. Although the assays and endpoints that are part of the OECD conceptual framework cover important aspects of disruption of the endocrine system and are generally recognised to be useful for the identification of chemicals with ED properties, it is also widely acknowledged that the OECD framework has considerable deficiencies with respect to relevant ED endpoints (see section 3.1). When compared to

the assays, model systems and endpoints that are available as scientific tools in the field of endocrine disrupter research, and which could potentially be further refined and taken forward for testing and screening, it becomes obvious that the OECD conceptual framework only deals with a small segment of these endocrine disrupting effects, mainly limited to the disruption of the estrogen, (anti)androgen, thyroid and steroidogenesis pathways (see section 4.7).

It should be borne in mind that these limitations are **particularly virulent in the arena of ecotoxicology**. For many taxa, internationally agreed guidelines and test methods for ED properties are missing.

7.1.1 IMPLEMENTATION OF TESTS FOR ED PROPERTIES IN PPPR

In view of the hazard-based exclusion criterion for approval of substances with ED properties that is enshrined in PPPR, it is necessary to improve the ability of tests to adequately demonstrate ED properties. This means that an update of Commission Regulations on data requirements for active substances (544/2011) and products (545/2011) is essential. To be fit for purpose, it must include further tests for ED properties.

7.1.1.1 General principles

Although important implications of endocrine disrupter definitions for the regulation of substances with ED properties are currently left open, it is widely accepted that testing in whole organisms is required to identify an endocrine disrupter. This has to encompass sufficiently sensitive endpoints of toxicological relevance that allow judgements in terms of adversity. To avoid that relevant effects are overlooked, the administration of test compounds has to cover recognised periods of sensitivity, windows of susceptibility and life stages.

To meet the requirements of the hazard-based exclusion criteria, together with the principles deriving from the definition for endocrine disrupters, two types of assays and tests must be incorporated, those that are able to demonstrate **adverse effects in whole organisms** which are considered to be relevant for humans and/or wildlife populations, and those capable of capturing an **endocrine mode of action**.

To realise the level of proof that is deemed necessary for the demonstration of adversity in human and mammalian toxicology, test outcomes and data from Level 5 assays of the OECD framework are essential and will have to be included in the mandatory data requirements for active substances in PPPR.

However, there is a fundamental principle at EU level that additional animal testing is not normally supported, without providing justifications. The justification for further animal use in ED testing flows directly from the WHO/IPCS definition of endocrine disrupters, the lack of information about correlations between the test outcome of OECD level 2 assays and assays at higher levels and the need to adminster chemicals during periods of heightened sensitivity which is not possible with in vitro assays.

Defining test requirements that ensure protection of wildlife from ED effects is more complicated, mainly because little consideration has been given to the question which tests can indicate adverse consequences at the population level by an ED relevant mode of action. Apart from OECD Level 5 assays, Level 3 and 4 assays for non-mammalian species may also be of value to derive indications of ED properties of chemicals on wildlife, on a case-by-case basis.

7.1.1.2 ED properties in human and mammalian toxicity testing

From the description of current testing requirements in section 5.3.1 it is obvious that important endpoints and tests required for the demonstration of ED properties of a chemical are currently not included. In accordance with the general principles laid out above, assays capturing adversity and endocrine modes of action need to be incorporated.

A minimum requirement and one that can be achieved in the near future is the addition of endpoints sensitive to endocrine disruption in studies of reproductive toxicity. This includes nipple retention, anogenital distance at birth and measurement of thyroid hormones. Further, in the context of PPPR, inclusion of developmental neurotoxicity and immunotoxicity are of particular relevance. This can be realised by incorporating these measurements in the currently prescribed two-generation reproduction toxicity study (TG 416) or by adopting the extended one-generation study (TG 443) which includes these endpoints.

OECD level 2 assays, such as those for estrogen or androgen receptor binding affinity, estrogen receptor transcriptional activation (TG 455), androgen and thyroid transcriptional activation, steroidogenesis *in vitro* (TG 456), MCF-7 cell proliferation and other assays, as appropriate, will provide valuable indications of endocrine modes of action. Such assays will also have to be included in the testing requirements.

In the future, testing requirements should also include the tests and endpoints described in the OECD Detailed Review Paper on Novel Tests and Endpoints for endocrine disruption, as and when further validated assays become available.

The combined chronic toxicity/carcinogenicity test (TG 451-3) is deficient in identifying endocrine disrupters with carcinogenic potential by a hormonal mechanism. This is particularly relevant for cancers of reproductive organs and the mammary gland. Demonstration of such properties will necessitate the development of new testing regimens, with new animal strains.

Because of the lack of information about correlations of test outcomes from Level 2 assays with those at Level 5 (see section 3.1.3), Level 2 assay outcomes cannot presently be used to filter out substances for which further testing at Level 5 could be waived.

It should be a matter for further analysis and debate whether there is value in including Level 3 assays (uterotrophic and Hershberger) in testing requirements for ED properties. These assays are useful for capturing toxicokinetic influences that might modulate or abrogate the effects seen in *in vitro* assays. On the other hand, it can be argued that the influence of toxicokinetics will be covered with Level 5 assays, making Level 3 test outcomes somewhat less informative.

The Oral 28-day study (TG 407) and the Oral 90-day study (TG 408) are both part of the current information requirements under PPPR, but their usefulness in detecting chemicals with ED

properties is limited, due to the omission of exposure during relevant windows of susceptibility, and to the recognised lack of sensitivity of the measured endpoints.

7.1.1.3 Ecotoxicology and non-mammalian toxicology

Under the current testing requirements set out in Regulations 544/2011 and 545/2011 there is only scope for the detection of ED effects on birds and possibly fish, but not on other non-mammalian taxa. It will need considerable efforts to work out the details of test requirements for the demonstration of ED properties of chemicals in wildlife. The issue is complicated by the fact that test guidelines have been worked out for only a few of the tests currently under discussion within the OECD framework.

With the exception of the sediment water chironomid life cycle toxicity test (TG 233) and the Daphnia reproduction test (TG 211) practically no other test guideline for the Level 5 assays that are currently under discussion is available. At Level 4 there is only the avian reproduction assay (TG 206), the chironomid toxicity test (TG 218-219) and the fish sexual development test (TG 234.) No other test guidelines for Level 4 assays for non-mammalian testing have been worked out, although quite a few are under discussion.

In view of this situation, an update of the information requirements for PPPR will have to rely also on the currently available OECD Level 3 assays for non-mammalian toxicology. To realise the demonstration of wildlife ED effects, inclusion of all validated test guidelines for non-mammalian species at Levels 3, 4 and 5 will have to be considered.

7.1.2 IMPLEMENTATION OF TESTS FOR ED PROPERTIES IN REACH

The relevant regulation defining testing requirements under REACH is the Test Methods Regulation (440/2008). As described in section 5.2, the testing requirements under REACH are differentiated according to supply tonnage. Of particular relevance to identifying ED properties of substances are the testing requirements laid down for reproductive and developmental toxicants.

7.1.2.1 General principles

The general principles elaborated above for PPPR relevant substances (see section 7.1.1.1) also apply for chemicals under REACH. Answers to the question which assays should be implemented for the various tonnage bands in REACH require quite complicated analyses which are beyond the scope of this report. Only some quite general considerations will be elaborated, with an emphasis on high production volume chemicals.

7.1.2.2 Endpoints relevant to ED properties in humans and mammalians

As detailed in section 5.2.1.4., the reproductive and developmental toxicity study currently conducted for chemicals in the highest tonnage band (> 1000 t/year), the two-generation reproduction study (TG 416) does not include the most sensitive endocrine related endpoints. Further, if a reproductive effect had already been demonstrated in another assay, such as repeated

dose toxicity studies (TG 407 and TG 408), this may then be used as the basis for the risk assessment and the requirement for a multigeneration reproduction study may be waived. To improve the ability of these tests to identify chemicals with ED properties, it is essential to include the endpoints discussed above for section 7.1.1.2., especially nipple retention, changes in anogenital distance, measurements relevant for thyroid disrupting effects, neurodevelopment and immunotoxicity.

The demonstration of ED-related modes of action will require implementation of OECD Level 2 assays.

It is worth serious consideration whether OECD Level 2 assays should also be required for lower tonnage bands as triggers to signal needs for further testing.

7.1.2.3 Endpoints relevant for wildlife ED properties

ECHA guidance on information requirements for environmental toxicity testing clearly states that "*there is no requirement set out in REACH Annexes VII to X to provide information on the endocrine activity of a substance or on a substance's reproductive or specific developmental toxicity in aquatic organisms".* Nevertheless, ECHA provides guidance on the evaluation of existing data for potential endocrine activity of a chemical, or long-term adverse effects on development and/or reproduction in aquatic organisms.

The recommended integrated assessment of available information (see Table 16 in section 5.2.1.5) will be useful as a basis for defining further tests. The assays suggested above as appropriate in the PPPR context are also of merit for identifying ED wildlife properties under REACH. However, it will require detailed analyses to work out which assay should be required in the various tonnage bands.

7.2 DECISION CRITERIA FOR THE REGULATION OF SUBSTANCES WITH ED PROPERTIES

Well-defined criteria are required for the translation of test outcomes into decisions about whether a substance should be regarded as possessing ED properties. In the context of the PPPR such decisions concern the question whether an active substance or formulation should be refused approval. Under REACH the question is which criteria should be met to designate a substance as being of concern equivalent to carcinogens, mutagens and reproductive toxicants, PBT and vPvB, and thus requiring authorisation.

Several EU member states have elaborated proposals for such purposes and these have been summarised in chapter 6 of this report.

This section will provide an overview of decision criteria, and will map out options for dealing with endocrine disrupters in the regulatory context of PPPR and REACH.

7.2.1 DECISION CRITERIA

Regulatory decisions about whether or not a chemical should be deemed an endocrine disrupter for regulatory purposes have to apply a series of criteria which derive from toxicological reasoning. These criteria should meet certain formal demands, as follows:

- **Consistency across regulations** – As much as possible, decision criteria should be consistent across the relevant sectorial regulations. It would lead to internal contradictions if the exclusion criteria under PPPR would be substantially different from those that merit designation of a chemical as a REACH Art 57 (f) substance.
- **Separation of hazard assessment from risk assessment** – This distinction is explicitly enshrined in the PPPR; only hazard-based criteria should be used to decide whether a substance is an endocrine disrupter. In any case, risk assessment is a step that can only be conducted subsequent to hazard assessment, another reason arguing for the separation of the two activities.
- **Separation of scientific considerations from socio-economic reasoning** – In the interest of protection of the public and the environment from possible harm, regulatory decisions about the hazardous properties of a chemical should not be clouded at the earliest stages by considerations of the possible socio-economic impact of these decisions. For example it is problematic when the stringency of decision criteria is increased unduly in view of the commercial implications of regulatory decisions. This is directly derived from the two points above, consistency across regulation regardless of the potential use or economical benefit of a substance, and a clear separation between hazard and risk assessment.
- **Transparency** – Classification criteria should be transparent, e.g. by following explicitly formulated weight-of-evidence evaluations.

Regulatory decisions as to whether test outcomes warrant designating a substance as an endocrine disrupter derive from toxicological considerations based on the following scientific criteria:

- **Adversity** – The effect should be adverse, according to the principles elaborated by WHO (see section 2.2.1). Certain assay requirements flow from this criterion. For all practical purposes, tests in intact organisms are required.
- **Mode of action** – An endocrine mode of action should be operating for the effect under consideration. This criterion is related to the additional criterion of specificity (see below).
- **Potency** – The potency of a substance in producing ED effects may inform weight-of-evidence considerations. Potency has also been proposed as the basis for hazard-based cut-off criteria (Joint DE-UK proposal, see 6.4). In this case, a continuous variable (potency) is in effect turned quantal (endocrine disrupter in the regulatory sense: yes or no).
- **Lead toxicity** – The observation that ED effects occur at lower doses than other toxic effects may play a role in weight-of-evidence considerations.
- **Specificity** – This can have three different meanings: Specificity in the sense of lead toxicity (see above), specificity in terms of an ED effect that manifests itself as a consequence of an ED-related mode of action, and not indirectly as a result of other systemic toxicity. In the context of ED effects on wildlife it can denote effects specific only for some species.

- **Severity** – Some ED effects can be judged to be more severe than others. Severity is important to evaluate whether the substance in question merits concerns equivalent to CMR. This criterion may be linked to that of irreversibility of effect (see below).
- **Irreversibility** – Certain ED effects are irreversible, and this can be used to argue levels of concern equivalent to other severe toxic outcomes such as CMR.
- **Relevance** – The relevance of an effect for human health is of importance. Similar considerations apply when judging an ecotoxicological test outcome in terms of its importance for deriving indications of adverse effects on wildlife populations.

7.2.2 DECISION TREES

The above criteria can be used to construct decision trees, as for example suggested by the joint DE-UK position paper (see section 6.4). Such decision trees array the above criteria in a particular order. The following stages can be distinguished:

Stage 1: Consideration of the quality of ED-related effects – adversity and mode of action

At this entry stage, test outcomes have to be evaluated in terms of evidence for ED properties (mode of action) and whether these properties give rise to adverse effects. The final outcome of the decision tree is influenced by the sequence in which these two decision criteria are applied. For example, if tests reveal clear evidence of an endocrine mode of action, but ambivalent results about adversity of effect, then a substance will not be pursued further in the decision tree, if adversity is the first criterion to be considered. However, tests at a later stage may reveal adverse effects, but by then the substance has then been "filtered out" in the decision tree. The converse is true when mode of action considerations form the entry point of the decision tree.

The case of paracetamol may give an example to illustrate the point. There is good evidence that paracetamol can induce effects indicative of androgen insufficiency in male offspring of dosed rats (changes in anogenital distance, retained nipples), but the mode of action that operates for these effects is less clear (see Annex 1, 3.5). If mode of action is the first decision criterion, then paracetamol would not be deemed an endocrine disrupter for regulatory purposes, although an endocrine mechanism is in fact operating, but not recognised.

To avoid decision dilemmas of this kind, it is important to **consider adversity and mode of action in parallel**, and not in sequence. This will yield a matrix that allows the differentiation of chemicals into endocrine disrupters, with strong evidence (good evidence for ED mode of action, and for adversity), with weaker evidence (suspected ED mode of action, and suspected adversity) and those where the evidence is mixed (either strong ED mode of action with suspected adversity, or vice versa). At this stage, the consideration of the criteria of mode of action and adversity can be decisive, and substances judged to be without adverse effects, and not following an ED mode of action may be regarded as not to be an endocrine disrupter for regulatory purposes.

In the interest of keeping the decision making transparent, it is important to apply **weight-of-evidence approaches**, for judging the quality of the studies, adversity and mode of action (see section 4.1). However, all evidence should be considered, without disregarding studies that have not employed guideline studies or have not followed GLP standards. Weight-of-evidence approaches for

evaluating adversity and mode of action for endocrine disrupters are not currently available and will need to be developed to make this stage viable.

At the end of stage 1, substances with weak evidence for the induction of ED-related effects **and** no indications for ED-related modes of action (upper left quadrant in **Figure 4**) are filtered out and not taken forward to the subsequent stages. Chemicals with no evidence for adverse effects, but indications of ED-related modes of action, and conversely, substances where evidence for adversity exists, but indications for ED-related modes of action are lacking, should be taken forward to the next stages of the decision tree.

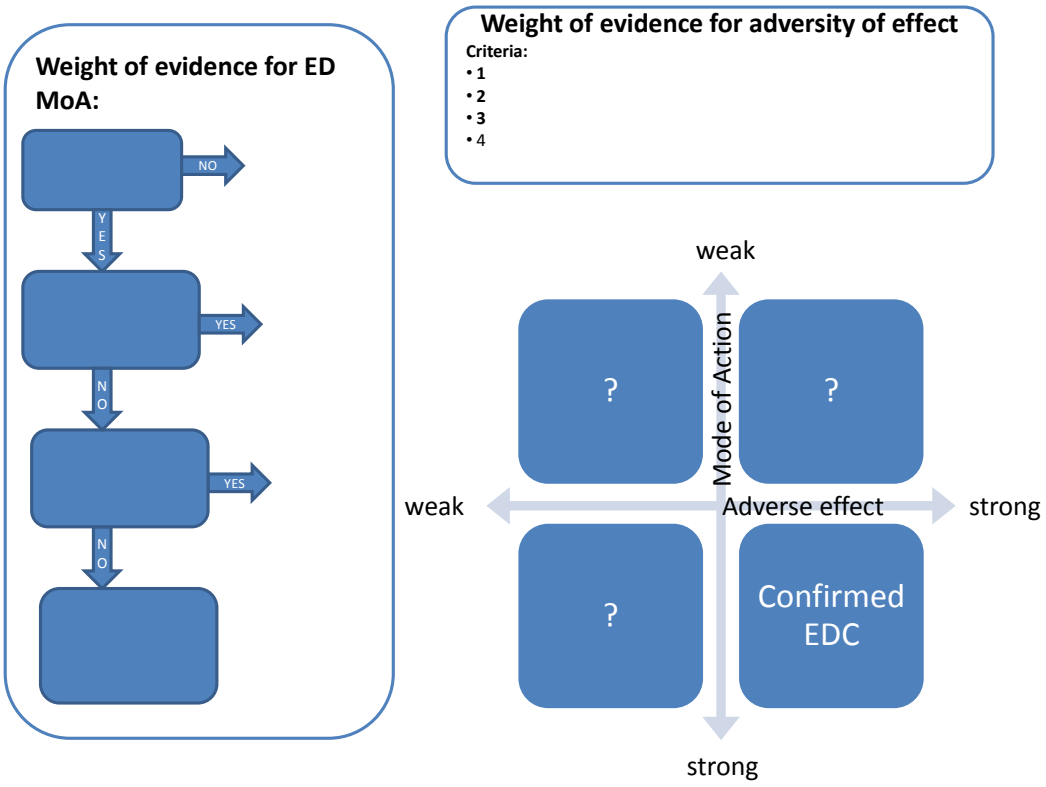

Figure 4: Evaluating evidence for adversity and mode of action together at Stage 1 of the proposed decision tree.

Stage 2: Evaluating human and wildlife relevance

At this stage, an analysis of the relevance of the observed effects for humans and wildlife is conducted, again with the help of weight-of-evidence approaches. These will have to be developed for human and wildlife relevance and are not currently available. All toxicological data should be taken into account. In the absence of appropriate scientific data, relevance should be assumed by default. However, a clear lack of human or ecotoxicological relevance, well supported by scientific data, should be decisive at this stage, with substances leaving the decision process.

Stage 3: Toxicological evaluation

The next stage comprises of a toxicological evaluation of test data, in terms of the criteria “potency”, “lead-toxicity”, “severity”, “specificity” and “irreversibility”. However, none of these criteria on their own should be decisive at this stage, i.e. no test substance should exit the decision tree and be regarded as not relevant for regulatory purposes. Instead, all these criteria should inform the final decision making on the basis of a weight-of-evidence approach, where **all the evidence is considered**. Again, such approaches have to be developed for endocrine disrupters.

Stage 4: Final decision, classification and categorisation

A final decision is made as to whether the substance in question should be regarded as an endocrine disrupter in the regulatory sense. For PPPR this means a decision as to whether approval should be withheld, according to the hazard based cut-off criterion. For REACH the decision will be whether the test chemical requires authorisation. The appropriate weight-of-evidence procedures for this stage have to be worked out.

7.2.3 POTENCY AS A FINAL DECISION CRITERION?

As detailed in section 6, the use of potency-based STOT-RE Cat 1 trigger values as cut-off criteria for endocrine disrupters of regulatory concern is highly controversial. It is supported by Germany and the UK, but categorically rejected by Denmark, France and several NGOs, including Chem Trust, WWF and PAN.

It has been argued that potency-based trigger values have the advantage of operationalising a concern equivalent to CMR – only very potent ED substances warrant this level of concern. The values have also been justified by referring to their wide acceptance in the context of CLP.

On the other hand, there are several problematic issues with potency-based trigger values when they are applied in a decisive way at the last stage of the decision tree, as suggested by Germany and the UK:

- Scientifically, it is impossible to draw a borderline for potency in isolation, without considering exposure. As such, solely potency-based trigger values will always be arbitrary.
- With many endocrine disrupters, the period of exposure, especially during critical windows of susceptibility is decisive for the induction of irreversible effects. The potency of a compound, although important, may not necessarily be the sole deciding factor. Potency-based cut-offs, when handled in a decisive way, do not take account of this feature of chemicals with ED properties. Potency-based cut-off values, especially if they are too restrictive, bear the danger that chemicals capable of inducing irreversible effects during critical windows of exposure remain unregulated.
- The introduction of potency-based trigger values violates the demand for consistency across various pieces of EU regulation. The potency-based trigger values derived from CLP were intended for purposes related to labeling and packaging for substances with target organ toxicity after repeat exposure. They are not suitable for substances of concern equivalent to CMR. Indeed, potency-based cut-offs are not in use for reproductive and developmental toxicants, due to the high concern about their effects.

- The submission of the Danish EPA (2011)[83] on criteria for endocrine disrupters contains an examination of potency estimates for selected endocrine disrupting chemicals to which the potency-based cut-off values proposed by Germany and the UK have been applied. As this assessment has made apparent, only very few of the chemicals regarded by many as endocrine disrupters, including phthalates, PCBs, and carboximide pesticides would be regarded as "endocrine disrupters" in a "regulatory sense" if these potency-based cut-off values were applied. The tables are reproduced below.

Updated version (of 17 May 2011) of the Annex 2 mentioned in relation to comments to 4: BAuA document on REACH, human health criteria, "Human health criteria for endocrine disruption (ED) according to Art. 57 (f) of the REACH regulation: German approach to the identification of EN substances as SVHC", 20th October 2010

Annex 2. Overview of effect levels determined for known endocrine disruptors (provided by DTU Food, the National Food Institute), updated 17 May 2011.

NOAELs and LOAELs for some substances with endocrine disrupting properties and adverse effects, and comparison of LOAELs with STOT RE guidance values for 90-day studies, i.e. 10 mg/kg bw/d for Category 1 and 100 mg/kg bw/d for Category 2. Please note that the LOAELs in many cases are based on decreased AGD and/or increased nipple retention and these endpoints are not included in the current OECD Guidelines.

Substance	NOAEL mg/kg bw/day	LOAEL mg/kg bw/day	Adverse effect(s) at LOAELs	LOAELs below 10 mg/kg bw/d?	LOAELs below 100 mg/kg bw/d?	References
DEHP	3 100 5	10 300	↓ AGD, ↑ Nipple retention, rat ↓ Testosterone GD 18, rat *Reproduktion (germ cell depletion, ↓ testis weight), developmental tox, rat*	Maybe?	Maybe?	Christiansen et al (2010) (27) Howdeshell et al 2008 (1) *Wolfe and Leyton, 2003 (*)EU RAR, EFSA*
DiNP	750 300 -	900 600 *750*	↓ AGD, rat ↑ Nipple retention, rat ↑ Nipple retention, rat	No	No	Boberg et al (2010) (28) Boberg et al (2010) (28) *Gray et al 2000 (2)* Exxon 1996 (*)
DnBP	- 250 50 10 100 -	250 500 250 50 300 (52) **2**	↓ AGD, rat ↓ AGD, rat ↓ AGD, rat ↓ Testosterone GD 19, rat ↓ Testosterone GD 18, rat Embryotoxicity, rat *Germ cell development, mammary gland changes,*	No	No	Ema & Miyawaki 2001 (3) Jiang 2007 (4) Zhang 2004 (5) Lehmann et al 2004(6) Howdeshell et al 2008 (1) Wine et al 1997 (7) *Lee 2004 (8)*

[83] DK EPA (Danish Ministry of the Environment, Environmental Protection Agency) (2011a) Establishment of criteria for endocrine disruptors and options for regulation. Dated 17 May 2011. Published as Annex 2 to chapter 4, comments to the BAuA document.
http://www.mst.dk/NR/rdonlyres/54DB4583-B01D-45D6-AA99-28ED75A5C0E4/127098/DKEDcriteria110517_finalcorr1.pdf

Substance	NOAEL mg/kg bw/day	LOAEL mg/kg bw/day	Adverse effect(s) at LOAELs	LOAELs below 10 mg/kg bw/d?	LOAELs below 100 mg/kg bw/d?	References
			rat			
DiBP	**125** *100*	250 300	↓ AGD, ↑ Nipple retention, rat ↓ *Testosterone GD 18, rat*	No	No	Sallenfait et al 2008 (9) *Howdeshell et al 2008 (1)*
BBP	***50*** 167 100 100 185 182	250 250 500 300 375 910	↓ *AGD, rat* ↓ AGD (GD 21), rat ↓ AGD, rat ↓ Testosterone GD 18, rat Developmental toxicity, mice Developmental toxicity, rat	No	No	*Tyl et al 2004 (10)*Ema et al 2003 (11) Nagao et al 2000 (12) Howdeshell et al 2008 (1) Ema et al 1990 (19) Price et al 1990 (26)
Prochloraz	5 ***3,7***	10 13	↑ Nipple retention, rat *Reproductive toxicity, rat*	Maybe	Yes	Christiansen et al (2009)(29) *Cozens et al 1982 (*)*
Epoxiconazol	**2,3**	23	Rat, 2-gen study, repro	No	Yes	Hellwig & Hildebrand 1992 (*)
Linuron	***0,8-1*** 10 25	 50	*Reproductive toxicity* Developmental, rabbit ↑ Nipple retention, rat	No	No	McKintyre et al 2000 (13)
Vinclozolin	- **5** *4* 4,9	5 10 -	↑ Nipple retention, rat ↓ AGD, rat 2 gen, reproductive toxicity, rat Reproductive toxicity, rat	Yes	Yes	Hass et al 2007 (14) Hass et al 2007 (14) *Hellwig et al 1994, BASF (*)* Hellwig et al 1990, BASF (*)
Procymidon	10 12,5 12,5 2,5	25 37,5 125 12,5	↑ Nipple retention, ↓ AGD, rat ↓ AGD, hypospadia, rat ↓ AGD, hypospadia, rat ↓ *AGD*, hypospadia, testis effekt, rotte	No	Yes	Hass et al 2007 (14) Wickramaratne et al 1998 (*) Hoberman et al 1992 (*) *EFSA scientific report 2009*
PCB's Arochlor 1254 Arochlor	- - - -	30 0,05 0,1 0,01	↓ AGD, ↓ Testosterone (↑ AGD, ↑ prostate weight, mice) ↓ AGD, ↓ organ weights, ↓	Yes	Yes	Lilienthal 2006 (15) Gupta 2000 (16) Faqi 1998 (17) Faqi 1998 (17)

7.2.4 OPTIONS FOR DEALING WITH ENDOCRINE DISRUPTERS IN VARIOUS REGULATORY CLASSIFICATIONS

It is possible to implement criteria for endocrine disrupters in the context of relevant pieces of EU regulations without developing wyas of handoling these chemicals in terms of regulatory classifications. Nevertheless, some options for dealing with endocrine disrupters in regulatory classifications are discussed here. EDCs could be subsumed under the class of STOT-RE in CLP. Alternatively, they could be dealt with as part of CMR substances. Finally, a separate class "EDC" could be created.

In this section, the advantages and disadvantages of these three options will be discussed.

7.2.4.1 Endocrine disrupters as part of STOT-RE

Dealing with endocrine disrupters as part of STOT-RE has a degree of plausibility, considering that many ED effects manifest themselves at the level of target organs.

On the other hand, STOT-RE endpoints typically focus on tests and assays that use adult animals, and do not cover the windows of susceptibility in development that are so important in identifying the effects of endocrine disrupters. It can therefore be argued that the implied parallel with target organ toxicants is not valid, particularly because STOT-RE substances do not pose concerns equivalent to CMRs. Furthermore, STOT-RE does not deal with wildlife effects.

Subsumption under STOT-RE will trigger major changes of the CLP regulation which will take very long times to implement. There is an agreement at EU level that after adoption of the CLP regulation new criteria will only be introduced if there is a request from the UN Global Harmonised System (GHS). Introducing changes in criteria via UN GHS takes years to accomplish.

7.2.4.2 Endocrine disrupters subsumed under CMR

With the argument that endocrine disruption represents multiple mechanisms that manifests themselves mostly in terms of carcinogenicity and reproductive toxicity, endocrine disrupting chemicals could be dealt with under CMR. This would seem appropriate and fitting, particularly in view of precedents with some endocrine disrupters that are already classified as reproductive toxicants. A requirement for realising this option is that the relevant test methods are implemented as part of the testing requirements for carcinogens and reproductive toxicants.

The drawback of this option is in the possibility that some endocrine pathways may trigger effects that lie outside the confines of CMR. Indeed it is likely that future research will substantiate this prospect. A second disadvantage is that, like STOT-RE, CMR does not incorporate wildlife effects.

7.2.4.3 Endocrine disrupters as a separate class "ED"

Finally, endocrine disrupters could be separated from CMR and STOT-RE and subsumed into a separate class "ED". The advantages of this option are manifold: It would afford sufficient flexibility to accommodate progress in scientific knowledge and lend the development of appropriate assays a fitting framework. An added benefit is in the fact that human and wildlife effects can be dealt with in one and the same class in a coherent fashion.

The disadvantage is that this option may require extremely time consuming amendments of the CLP via UN GHS (see above).

7.2.5 CATEGORISATIONS AND STIMULI FOR REGULATORY ACTION AND FURTHER TESTING

Another area of major controversy concerns the question which evidence should be required to refuse approval for a substance due to endocrine disrupting effects. It is the explicit intention of the joint DE-UK proposal to regulate only substances with ED properties that pose very high concern, and to leave substances that fall below the chosen criteria unregulated.

In contrast, the Danish proposal takes a less restrictive and more precautious stance, by introducing subdivisions of categories that reflect the strength of the available evidence ("confirmed", suspected", or "indicated", see section 6.6.2). To create incentives for generating further data, certain categories in the Danish proposal stipulate further testing.

In view of the large gaps left by currently available test methods, and the concerns about inadequate protection of humans and wildlife that may stem from these gaps, the Danish approach appears to be advantageous. Indeed, if inadequately tested substances remain in categories that do not trigger regulatory action, that situation will create strong disincentives for further testing.

However, the general problem lies in the qualitative lack of validated test methods that are capable of creating a higher degree of protection by capturing wider aspects of endocrine disruption. The demand for further testing for substances classed in lower categories may therefore provoke the questions as to the gains this might bring, when further testing is limited to validated test methods.

The dilemma will be that validated test methods that could be used for further testing are simply not available. In the short term, the dilemma can only be resolved by extending the testing to methods that are not yet validated and internationally recognised. In the longer run, it can only be dealt with through continued and speedy efforts of developing, validating and implementing better test methods.

7.3 RECOMMENDATIONS

On the basis of the above considerations, the following recommendations can be made:

- Implement recently updated or enhanced validated and internationally recognised test methods in the testing and information requirements for PPPR and REACH,
- Develop further guidance documents for the interpretation of test data,
- Consider the creation of a separate regulatory class “Endocrine Disrupter” (ED),
- Develop weight-of-evidence procedures that deal with the available evidence by weighing the criteria “adversity” and “mode of action” in parallel, but not by applying these criteria sequentially to exclude substances from the assessment,
- Consider potency, together with other criteria such as lead toxicity, specificity, severity and irreversibility in a weight-of-evidence approach. Abandon “potency” as a rigid cut-off criterion for endocrine disrupters of regulatory concern, for lack of prospect of reaching a consensus by purely scientific criteria,
- Create regulatory categories that stimulate the generation of the necessary data, including test methods that are not validated, beyond the OECD Conceptual Framework.

8 RESEARCH AND DEVELOPMENT NEEDS

The Summary of the State of the Science on Endocrine Disrupters (Annex 1) has exposed knowledge gaps that considerably hamper progress with the risk assessment and regulation of EDCs. These gaps exist despite years of extensive research efforts. They are explained largely by a latent conflict between the drive for novelty by scientists and the needs of the chemical regulatory arena which emphasise more routine tasks, not always congruent with the interests of scientists.

It is beyond the scope of this report to provide an in-depth analysis of research needs. This will have to be achieved as part of a **thorough gap analysis**. However, elements of future research needs can be distinguished in three main areas:

- Exposure assessment and identification of substances with ED properties
- Assay development
- Human epidemiology

8.1 EXPOSURE ASSESSMENT AND IDENTIFICATION OF SUBSTANCES WITH ED PROPERTIES

Our assessment of the endocrine disrupter science has shown that ED research still focuses unduly on a relatively small subset of chemicals. These substances are the topic of scientific investigations, not necessarily because they are priority pollutants, but simply because they are already well researched. Too little work has been conducted to systematically screen chemicals for their ED properties, and to some degree this effort is hampered by the lack of high through-put assays that capture ED effects beyond the usual modalities of (anti)estrogenicity, (anti)androgenicity and thyroid disrupting properties (see below).

The result is that the full spectrum of chemicals that potentially contribute to endocrine-related diseases is far from known.

The gap could be closed by adopting unbiased exposure assessment strategies that search for unknowns by mobilizing recent advances in chemical analytical technology and by exploiting methods that use ED mode-of-action screens to interrogate extracts from complex environmental media and human tissues (effect-directed analysis). Such strategies have the potential of pinpointing previously unrecognised endocrine disrupters and are a rich topic for research and development.

8.2 ASSAY DEVELOPMENT

Assay development is an arduous task which many scientists shy away from, due to uncertain rewards in terms of novelty which may impede the drive for high impact papers. Nevertheless, assay development will be crucial if the risks potentially deriving from EDCs are to be properly assessed. This conflict between the needs of scientists and the regulatory arena cannot be resolved without directed research and the targeted allocation of resources and funding.

There is a wide variety of assays, models and tools that scientist have developed for the study of ED related modes of action and mechanisms. Many of these could be taken forward and developed into validated assays to cover aspects of the endocrine system that are outside current testing strategies. This is essential to reduce the uncertainty in assessing human health and wildlife risks that stems from the current gaps in the available tests.

Assays for many wildlife phyla and taxa are currently not developed, and this is a rich field for future research and development.

While many existing “scientific tools” could be refined into assays for use in the regulatory arena, there are aspects of endocrine disruption that are currently not accessible to systematic investigation, because suitable models are missing altogether. This is especially relevant in the areas of metabolic syndrome, obesity, neuro-endocrine effects. Concerted research and development efforts are needed to fill these gaps. In these areas there is currently no alternative to relying on epidemiology as a means for hazard and risk characterization. Because epidemiology is not the most sensitive tool for hazard and risk identification, and because risks can only be identified after the event, this introduced considerable uncertainties, with a strong likelihood of overlooking effects.

8.3 HUMAN EPIDEMIOLOGY

In human epidemiology there are considerable difficulties in finding ways of recognising the health risks that may stem from endocrine disrupters. Complications arise mainly from the time lag between disease causation and the diagnosis of health effects, the absence of methodologies for dealing with exposures to multiple chemicals in epidemiology and the lack of information about the full spectrum of chemicals that might contribute to risks.

Dealing with the problem of time lag between exposure and effect will require resources to set up new cohorts, with carefully planned measurement strategies and clearly worked-out hypotheses. The tissues collected in many existing cohorts are not geared towards ascertaining exposures to polar chemicals, and towards dealing with exposures during critical life stages. Whole blood biobanks are ideal for the study of highly lipophilic pollutants, but are not suitable for analysing polar pollutants, because of loss through enzymatic digestion upon thawing.

A dialogue is beginning between experimental mixture toxicologists and human epidemiologists to find ways of dealing with combined exposures in human health studies. These efforts should be nurtured and pursued further in the future. The topic of endocrine disrupter research is a fertile ground for developing novel concepts in epidemiology.

9 ANNEXES

9.1 SUMMARY OF THE STATE OF THE SCIENCE ON ENDOCRINE DISRUPTION (ANNEX 1)

9.2 SUMMARY OF EXPERT CONSULTATIONS ON APPROACHES TO THE REGULATORY ASSESSMENT OF ENDOCRINE DISRUPTERS (ANNEX 2)

9.3 COMPARATIVE ANALYSIS OF ENDPOINTS AND ASSAYS BY HUMAN HEALTH AND WILDLIFE ENDPOINT (ANNEX 3)