

STATE OF THE ART ASSESSMENT OF ENDOCRINE DISRUPTERS

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

Annex 2

SUMMARY OF EXPERT CONSULTATIONS ON APPROACHES TO THE REGULATORY ASSESSMENT OF ENDOCRINE DISRUPTERS

Authors: Michael Faust, Andreas Kortenkamp

CONTENTS

1	Aims and approach	3
2	Overview on expert views and opinions.....	6
2.1	Definition of EDs – requirements for regulatory purposes.....	6
2.2	Identification of EDs – views about tests and testing strategies.....	13
2.3	Hazard based grouping and classification of EDs.....	19
2.4	Evidence based categorizations of EDs	27
2.5	Cumulative exposure and mixture effects of EDs	32
2.6	Regulatory assessment of EDs – What signifies an “equivalent level of concern”?	34
2.7	Pragmatic default assessment criteria for EDs	38
2.8	Complementarities of national approaches and Community approaches.....	40

1 AIMS AND APPROACH

As agreed at the kick-off meeting, informal interviews have been conducted with experts from the following 18 organisations:

EU Member States agencies

- 1 Denmark, Miljøstyrelsen (Environmental Protection Agency)
- 2 France, Agence Nationale de Sécurité Sanitaire (ANSES)
(French Agency for Food, Environmental and Occupational Health Safety)
- 3 Germany, Bundesinstitut für Risikobewertung (BfR) (Federal Institute for Risk Assessment)
- 4 Germany, Umweltbundesamt (UBA) (Federal Environment Agency)
- 5 Italy, Istituto Superiore di Sanita (ISS) (Superior Health Institute)
- 6 Netherlands, Rijksinstituut voor Volksgezondheid en Milieu (RIVM)
(National Institute for Public Health and the Environment)
- 7 Sweden, Kemikalieinspektionen (KEMI) (Swedish Chemical Agency)
- 8 United Kingdom, Department for Environment, Food and Rural Affairs (DEFRA)

Competing economies

- 9 Canada, Health Canada
- 10 Japan, National Institute of Health Sciences (NIHS)
- 11 Switzerland, commissioned representative in the OECD EDTA AG
- 12 USA, United States Environmental Protection Agency (US EPA)

Intergovernmental agencies

- 13 Nordic Co-ordination Group for the Development of Test Methods in Toxicology and Ecotoxicology (Nord-Utte)
- 14 Organisation for Economic Co-operation and Development (OECD)

Industry associations

- 15 The European Chemical Industry Council (CEFIC)
- 16 European Centre for Ecotoxicology and Toxicology of Chemicals (ECETOC)

Environmental charities

- 17 International Chemical Secretariat (ChemSec)
- 18 Chemicals, Health and Environment Monitoring Trust (CHEM Trust)

The aims of the interviews were

- to ensure that the report on the state-of-the-art assessment takes note of all relevant current, ongoing or planned activities and approaches in EU MS regarding identification, testing, assessment and regulatory management of EDs, and
- to ensure that all points considered critical and important by experts of MS authorities for a state-of-the-art assessment of EDs are properly reflected in the contractor's report to the Commission.

The interviews followed an interview guide that was prepared by the Contractor and 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. The guides have been documented as Annexes to the first interim report.

The interview guide covered the following eight topics:

1. Definition of EDs – requirements for regulatory purposes
2. Identification of EDs – views about testing strategies
3. Hazard based grouping and classification of EDs
 - MoAs
 - endpoints
 - potencies and / or specificities
4. Evidence based categorizations of EDs
 - guiding criteria
 - dealing with gaps and uncertainties
5. Cumulative exposure and mixture effects of EDs
 - options within existing legislation
 - new complementary approaches
6. Regulatory assessment of EDs – What signifies an “equivalent level of concern” ? (REACH)
7. Pragmatic default assessment criteria for EDs (PPP Regulation)
8. Complementarities of national approaches and Community approaches

The interviews had the character of informal consultations. It was not intended to receive official statements of the agencies represented by the interview partners. In this report, an overview of the outcome of the consultations is given in a way that is not attributable to individual interview partners.

The interviews were conducted during the period from April to October 2010.

The following overview on expert opinions is structured according to the list of topics given above. For every topic a brief outline of the background, the questions address, a summary of the answers, and a compilation of selected anonymised responses is provided that provides an overview on the spectrum of views and opinions received.

2 OVERVIEW ON EXPERT VIEWS AND OPINIONS

2.1 DEFINITION OF ENDOCRINE DISRUPTER (EDS) – REQUIREMENTS FOR REGULATORY PURPOSES

Background

There are several definitions for endocrine disruptors (ECETOC 2009¹, BfR 2010²). In the EU, the “Community strategy for endocrine disruptors” specifically refers to a “working” definition developed by WHO/IPCS:

“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” (COM(1999) 706 final, p.5-6)³.

Questions

- Does this *working* definition sufficiently serve regulatory purposes or are amendments necessary?
- Is this definition appropriate for the protection of both human health and the environment?

Experts from organisations other than EU MS authorities were additionally asked:

- Does your authority prefer an alternative definition, and if yes, which and why?

Answers

¹ ECETOC (European Centre for Ecotoxicology and Toxicology of Chemicals) (2009) Guidance on Identifying Endocrine Disrupting Effects. Technical Report No. 106, ECETOC, Brussels

² BfR (German Federal Institute for Risk Assessment) (2010) 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. January 15th, 2010, Federal Institute for Risk Assessment, Berlin, Germany

³ COM (Commission of the European Communities) (1999) Community Strategy for Endocrine Disruptors, a range of substances suspected of interfering with the hormone systems of humans and wildlife. Communication from the Commission to the Council and the European Parliament, COM(1999) 706 final, Brussels, 17.12.1999

The WHO/IPCS definition of EDs is generally acknowledged as one of the best. Some of the interviewees consider some of the other available definitions to be fairly equivalent, but do not see that any of them is preferable over the WHO/IPCS definition.

Part of the consulted experts is convinced that the WHO/IPCS definition of EDs is not only consensually acceptable but that it sufficiently serves all regulatory purposes. Some of them even warned against any amendments. For example:

At this stage of implementation of EU laws and regulations it would be inappropriate and unwise to change the definitions which have been thoroughly discussed at an international level over the course of many years.

Others, however, think that for regulatory purposes amendments should be introduced. A third group accepts the WHO/IPCS definition as a suitable working basis for all current regulatory tasks and activities, but sees a need for various complementary clarifications and specifications.

There is a general agreement, that the WHO/IPCS definition provides a top-level definition that is not only applicable to human health risk assessments but also suitable for ecotoxicological hazard and risk assessments. In detail, however, there is an obvious need for clarifications of the appropriate interpretation of the definition in an ecotoxicological context. Moreover, several experts state that regulatory use requires series of more detailed ecotoxicological criteria.

Major points of discussion that come out from the interviewees' response to the question of ED definition are the following:

- the concept of adversity,
- the proof of causality,
- the inclusion of mixtures in the definition, and
- the interpretation for ecotoxicological risk assessments.

Concept of adversity

Almost all interviewees stressed the point that adversity in an intact organism is the important concept in the WHO/IPCS definition. For example:

Adversity is a key criterion of the WHO/IPCS definition of endocrine disrupters. Adversity makes the difference between a mere endocrine modulator and an endocrine disrupter.

Several experts explained this in detail. For example:

Because the endocrine system is a communication system and changes in endocrine parameters are expected in response to changes in environmental conditions, not all changes will necessarily result in an adverse effect. For the same reason, any exposure to a xenobiotic will eventually result in a change in endocrine parameters. There were some early efforts to differentiate between primary, secondary and indirect endocrine toxicity, e.g. liver enzyme alteration would

affect circulating hormone levels. These concepts have kind of fallen by the wayside, but this is not that important for the definition of endocrine disrupters.

As a consequence, the WHO/IPCS definition of EDs crucially depends on the corresponding definition of the term “adverse health effects”. Several experts stated that it is acceptable, if used with reference to the definition of “adverse effect” that has been established in the IPCS risk assessment terminology: Any “change in the morphology, physiology, growth, development, reproduction or life span of an organism, system, or (sub)population that results in an impairment of functional capacity, an impairment of the capacity to compensate for additional stress, or an increase in susceptibility to the harmful effects of other environmental influences”⁴. Others pointed out that a slightly amended version of this WHO/IPCS definition of adversity provides an even more suitable working solution, as has been detailed in a BfR workshop report (BfR 2010)⁵.

Despite these existing definitions, several experts expressed that they are unsure how to handle the adversity criterion in practice. For example:

The concept of adversity in that definition should be elaborated. For example, should a compound that by itself does not induce hypospadias, but affects the expression of genes important for processes of male sexual differentiation, be considered as inducing adverse effects? By considering parameters related to gene expression, etc., the definition of adversity may change in the future.

Considerations of the adversity criterion lead to the question why adverse effects caused by EDs require a specific regulatory approach as compared to adverse effects caused by chemicals with other modes of action. One of the interviewees challenged this notion particularly clearly:

Regulation should be based on endpoints, not on mechanisms. The critical endpoint is reproductive toxicity as it can be detected in a two-generation animal study, irrespective of the modes and mechanisms of action by which this type of effect might be caused. There is no reason why EDs should be regulated separately and differently from other reproductive toxicants.

Another expert expressed a similar opinion:

Toxicity tests always investigate specific endpoints and e.g. dose/response relationships. The mechanism involved is, at the moment of testing, of secondary importance. For the evaluation of ED this means that classical toxicity tests provide sufficient evidence of harmful effects, irrespective of whether or not an effect is due to an endocrine related mechanism or another, not endocrine related mechanism.

One of those who represent an opposite standpoint expressed this as follows:

⁴ WHO (World Health Organisation) (2004) IPCS Risk Assessment Terminology. Geneva.

⁵ BfR (German Federal Institute for Risk Assessment) (2010) 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. January 15th, 2010, Federal Institute for Risk Assessment, Berlin, Germany

There is lack of clarity in some quarters as to why endocrine disrupters deserve special attention. Endocrine disrupters deserve attention because they can affect development during certain life stages, with delayed and irreversible effects on multiple target organs/tissues.

More detailed and expanded argumentations in favour of a special regulatory treatment of EDs were given by several other experts under the subsequent interview topics, in particular hazard and evidence based classification (2.3 and 2.4) and level of concern (2.6).

Proof of causality

The WHO/IPCS definition was provided as a working definition for an OECD workshop in September 2009. There were some discussions during the meeting as to whether the definition should be modified to allow for upstream events to be used instead of the adverse effect itself as an endpoint for the regulation of chemicals (see OECD EDTA AG 2009, Part I, p. 9)⁶. One of the interviewees pointed out:

This is important because it would allow newer toxicological methods to be used in hazard/risk assessment, however no general agreement could be reached.

Two other experts stressed this point in more detail and proposed a corresponding amendment of the WHO/IPCS definition:

The WHO/IPCS definition includes an establishment of causality. Establishment of such a causal link is in natural science normally done by showing that the independent variable (here extent of endocrine activity) correlates with the dependent variable (extent of adverse effects) backed up by various types of supporting evidence related to “modes of action” or “Adverse Outcome Pathways” and/or well established and/or plausible biochemical, physiological or biological theories/hypothesis.

For regulatory purposes, we would like to see the WHO/IPCS definition of endocrine disrupters amended in a way that allows to take upstream events into consideration and that does not in any case require the definite proof of a causal link between an ED mode of action and an adverse health outcome. A strong indication for such a causal link should be sufficient. To this end, the formulation “... and consequently causes adverse health effects in an intact organism,...” should be modified by insertion of a formulation such as “or is known to be causally linked to adverse health effects” after “causes” or “is likely to be associated to adverse health effects”. There should be no need for actually having seen e.g. reproductive malformations, but for regulatory action a strong indication for a causal link between an ED mode of action (MoA) and an adverse outcome should be sufficient.

Complementary to these proposals for amendments of the top-level ED definition, several interview partners stated that for regulatory purposes

⁶ OECD EDTA AG (Organisation for Economic Co-operation and Development, Advisory Group on Endocrine Disrupters Testing and Assessment (EDTA) of the Test Guidelines Programme) (2009) Workshop report on OECD countries activities regarding testing, assessment and management of endocrine disrupters, Part I. OECD document ENV/JM/TG/EDTA/M(2009)2/PART1.

the definition of EDs must be complemented by a corresponding definition of “potential” or “suspected EDs”.

The reasons for such statements were illustrated by examples like this:

Actual malformations are a relatively rare event that cannot often be detected in guideline studies. Decreased anogenital distance (AGD) in male rats, for instance, is a clear signal that something in the reproductive development of the individual has gone wrong, but a causal link to a specific ED MoA may be difficult to establish. There may also be cases where we face the reverse situation, i.e. where specific modes of ED action give strong reasons for concern about potential severe health effects although we have not yet actually seen them.

This aspect was addressed in more detail in the views about evidence-based categorizations under topic 2.4.

In the context of considerations about causality and adversity, two of the interviewees pointed to the fact, that relevant pieces of current EU legislation, such as REACH Article 57 (f), use the term “substances having endocrine disrupting properties” without any explicit definition. One of them concluded that this requires clarification. The other one stated:

It would be inappropriate to interpret the term “endocrine disrupting properties” as equivalent to the IPCS definition of EDCs. This is because it would be folly to require proof of the mechanism of action by which a chemical acted before it could be regulated.

Inclusion of mixtures in the definition

The WHO/IPCS definition encompasses both single substances and mixtures of substances. Some experts highlighted this as a positive feature that provides a basis for the assessment of cumulative risks from simultaneous exposure to different endocrine disruptors.

However, from one of the experts we heard a critical remark on this:

It is not useful to have the term “mixture” already included in the definition. The definition should strictly refer to substances. The inclusion of mixtures in the definition is only meaningful in the context of the regulation on classification and labelling, but not for REACH or other substance-oriented pieces of chemicals legislation.

Interpretation for ecotoxicological risk assessments

It is obvious, that the WHO/IPCS definition was originally developed in the context of considerations about human health. Nevertheless, the interviewees consensually considered the WHO/IPCS definition to be appropriate for the protection of both human health and the environment. The types of reasoning given were the following:

- *We interpret the term “health effects” in a wider sense, not restricted to human health, but including all types of organisms.*
- *The IPCS definition is also appropriate for ecotoxicological hazard assessments, if the term “health effects” is extended to populations as the protection target.*
- *The WHO/IPCS definition comprises both effects on individual organisms as well as on populations; hence it is applicable to both human health risk assessments as well as ecotoxicological risk assessments.*
- *The use of the term ‘(sub)population’ also makes clear that effects on wildlife and ecological targets are included in this definition.*

As consequence, the question arises how the concept of adversity should be interpreted in an ecotoxicological context. Two of the experts expressed views on this as follows:

- *Ecotoxicological assessments aim at the long-term conservation of species and populations, not at the protection of individual health. For this purpose, the WHO/IPCS definition of EDs is interpreted to cover substances that cause endocrine mediated population-relevant adverse effects.*
- *The definition refers to “adverse health effects in an intact organism or its progeny, or (sub) populations”. Many different definitions of adverse effects exist which may include increased susceptibility to naturally occurring stress factors. In general in ecotoxicology effects on survival, growth, development and reproduction as recorded in single species laboratory tests are regarded as ecologically relevant types of effects relevant for the maintenance of wild populations, which are the protection target. We agree in principle to the view that relevant effects may also be such that are more or less indirectly related to these types of effects, e.g. mating behaviour. The reference to subpopulations in this definition may then relate to also particular sensitive sub-populations of wild animals, if such are known to exist.*

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 with it should be understood in a very general way. This may have serious consequences in the context of pesticides and biocides regulation. In the interviews we heard the following opinions on this point:

- *Reference is made to “the endocrine system” and “organism” but it is not specified in the definition which types of organisms the definition relates to. Organisms also include plants, microbes and invertebrate taxa, but in general endocrine disruptors often implicitly refer to the endocrine system of vertebrates because the endocrine system of other types of organisms is in most cases not so well known or when known not regarded as a particular protection target as such.*
- *It should be noted here that the WHO/IPCS definition generally also covers chemicals which interfere with hormone systems in plants. As certain classes of pesticides are specifically designed to combat target plants (weeds) by this mode of action (so-called plant growth regulators), these deserve special consideration in risk assessment and management to prevent unacceptable side-*

effects to non-target plants. However, it is recognised that the debate on EDCs historically centred around effects on humans and animals. This is the context where the WHO/IPCS definition was developed.

- For regulatory purposes, uniformity of the ED definition and consistency of assessment and classification criteria across all the different pieces of chemicals legislation are absolutely crucial. Problems may arise where substances used as pesticides or biocides exert their wanted effects through an endocrine disrupting mechanism or mode of action (MoA). Appropriate regulatory approaches to a proper distinction between wanted endocrine disrupting effects in target pests and unwanted endocrine disrupting effects in non-target organisms have to be worked out.*

One of the experts pointed to the more problems that ecotoxicologist currently face in the regulatory risk assessment of EDs:

A regulatory context requires that the definition be associated with a series of more detailed criteria, that will be used to identify those substances worthy of regulatory action based on their ability to disrupt the endocrine system.

Under the current testing regimen, no single ecotoxicity test will provide the information necessary to define a chemical as an endocrine disrupter with regard to the above definition.

The three assays that have recently been adopted [by OECD] will determine mechanisms (or perhaps provide data on toxic end points associated with a particular mode of action) for the hypothalamus-pituitary-thyroid or the hypothalamus-pituitary gonadal axes, but will not provide information regarding adverse 'health' effects.

Of the current toxicity tests, only the fish early life cycle/full life cycle studies will give any adverse health effect data. The one-generation avian study will provide data on F0 reproductive effects and F1 developmental effects up to day 14 post-hatch. To use the above definition, a weight of evidence approach or extrapolation from the mammalian toxicity data would be required. Frog metamorphosis assays might also be useful.

These arguments directly lead to the second interview topic that dealt with tests and test strategies.

2.2 IDENTIFICATION OF EDs – VIEWS ABOUT TESTS AND TEST STRATEGIES

Background

As part of the medium term actions of the *Community strategy for endocrine disruptors*, the Commission, as well as some EU Member States, actively participates in the development and validation of agreed test methods and assessment guidelines by the OECD EDTA Advisory Group and its progenitor the OECD EDTA Task Force. The same applies to the competing economies that are included in this survey.

Questions

- Does the work on test methods and assessment guidelines performed by the OECD EDTA Task Force / Advisory Group satisfy regulatory needs or are additional efforts required? If yes, which ones?

Answers

All interviewees highly appreciated the work of the OECD EDTA Task Force / AG and stressed that it is valuable and important.

However, opinions were divided when it came to question as to whether the OECD work on test methods and guidelines satisfies regulatory needs. A minority answered this question with a clear yes. The others complained about gaps in available validated test guidelines and slow progress. Some further comments stressed the need for finding agreements on test strategies, assessment guidelines, and the use of results from OECD test under Community legislation.

Gaps in available test methods and needs for additional efforts

During the interviews, a variety of experts provided very detailed opinions about gaps in the OECD test methods for EDs and the consequences for future work. In the following 6 selected summaries of these opinions are given with the aim to cover the most important aspects from different angles.

Expert 1

The existing OECD test guidelines (TGs) serve the purpose of identifying EDs only partly. There are three major shortcomings:

- The existing TGs do not give a clear clue on whether a substance interferes with one of the four major pathways of ED action: E, A, T, or steroidogenesis.*

- ii. Existing screening assays for substances are too simplistic because they do not include the metabolites.
- iii. There is a need for assays that are able to give an indication for more complex modes of interaction of EDs with an endocrine axis, other than simple direct interaction with hormone receptor sites.
- iv. TG data (e.g. screening data) interpretation and their use in risk assessment and testing strategies are not clear.

Expert 2

The existing OECD test guidelines (TGs) serve the purpose of identifying EDs only partly. Major shortcomings are the following:

- i. The *in vitro* and the acute *in vivo* tests essentially only cover estrogenic, androgenic and thyroid mechanisms, while assays on the interference with other hormone axes, such as glucocorticoids, progesterone, or PPAR (peroxisome proliferator-activated receptors), are missing.
- ii. The existing long-term tests that cover all major life stages are largely restricted to endpoints of reproductive toxicity and inappropriate to detect other potential long-term effects of EDs on organ systems, such as effects on
 - a) brain development and adult function, regarding sexually dimorphic as well as cognitive and emotional functions (DNT, TG426, not appropriate because treatment stops before puberty),
 - b) non-sex hormone axes such as the adrenal axis (as discussed in the OECD Workshop in Copenhagen 2009),
 - c) lipid and glucose metabolism (energy metabolism, homeostasis),
 - d) ageing, and
 - e) epigenetic effects.
- iii. Existing reprotox guidelines still are less suited to detect alterations in functions; their emphasis is on persistent structural changes. Trivial to say that functions can be disturbed in the absence of structural changes.
- iv. The capacity to cope with environmental stress is of major importance for survival in the real world. However, there is no long-term (reprotox or other) test guideline that would allow to assess the “impairment of capacity to compensate for additional stress or increased susceptibility to the harmful effects of other environmental influences”, as required in the 2004 WHO/IPCS definition of adverse effects. The developmental immunotox (DIT) module of EOGRTS (Extended One Generation Test) would represent a first step in that direction, but limited to immune responses.

- v. *It cannot be safely assumed that LOAELs for all these aforementioned endpoints will consistently be similar to LOAELs for conventional reprotox endpoints. Conventional parameters such as histopathology and organ weights might be not sensitive enough to detect more subtle functional changes caused by EDs.*

Considering the logistics and the work load of existing or planned guidelines (TG416, EOGRTS), it becomes clear that it will be difficult to close all these gaps simply by adding more and more conventional test endpoints to the existing guidelines. As a matter of fact, the discussions in the expert group for the EOGRTS have shown that these long-term protocols are already at the limit of their capacity. That's why the developmental neurotox and developmental immunotox modules had to be reduced to the absolute minimum. It is therefore unclear how further endpoints of ED action, such as effects on brain functions or metabolism for instance, could be included. For reasons of animal welfare, it would also hardly be possible to conduct new long-term tests in addition to a two-generation assay or an EOGRTS for reproductive toxicity.

As a way out of this dilemma, we need new approaches, new endpoints based on modern techniques (molecular biology, genomics). However, before this makes sense, we need more research on causal links between molecular endpoints and adverse effects on the phenotype level. With the aim to encourage such a development, the OECD guideline system should become more flexible, allowing for swift revisions when new science-based endpoints appear to be mature for introduction.

Expert 3

Endpoints related to obesity, metabolic syndrome X and their cardiovascular effects would not be covered under the current scheme for testing. There are examples of compounds, such as tributyl tin, nicotine and nicotine control therapeutics, that are not estrogen or androgen receptor ligands and would not be detected by assays addressing those mechanisms of actions.

My personal view is that the next two decades will see a paradigm shift in toxicological testing, away from whole animal testing towards computational toxicology, read-across and in vitro bioassays, with animal tests simply used to validate that framework. Initiatives such as ToxCast in the US are still very experimental at present and many years away from replacing animal testing. Regulatory thinking is very rigid and pressure for change is likely to come from the industry as the burden of testing increases.

Expert 4

- *The existing OECD test guidelines are insufficient for the identification of EDs. A lot of work for improving the situation is already going on, but beyond that, additional efforts are required.*
- *In the 2-generation OECD assay*
 - *nipple retention (NR) is not assessed,*
 - *measurement of AGD is not obligatory (but only assessed in F2 as an endpoint triggered by changes in the sex ratio), and*

- *only one male per litter is examined for malformations of reproductive organs.*

The long-lasting debate about the necessary update of the 2-generation assays has led to the proposal for an extended 1-generation assay.

- *The proposed extended 1-generation OECD assay could be a clear improvement for the regulatory identification of EDs, but this will of course depend on the actual final version.*

Expert 5

Important gaps in the set of available test guidelines are the following:

- *Fish full life cycle assay: although crucial for ecotoxicological assessments, no OECD guideline is yet available. As a makeshift solution, UBA requires tests according to a US EPA guideline, supplemented by some endocrine endpoints.*
- *Snails: test is stuck in the pre-validation phase and thus not yet available.*
- *Birds: only a reproductive test but no life cycle assay is available.*

Expert 6

There is currently a data gap with regard to developmental effects in the F1 generation in birds (day 14 post hatch).

Attributing cause-effect relationships in an intact organism under the current testing regime is impossible without using a weight of evidence approach, as there are no tests/assays that evaluate both the mechanism and adverse effects in a single organism.

In total, the OECD test battery should be sensitive to a wide range of EDCs and non-EDCs, including inter alia (anti)oestrogens, (anti)androgens, aromatase inhibitors, juvenile hormone mimics, and ecdysone agonists and antagonists. However, although the vertebrates are well-represented in the battery, the only invertebrate representatives to date are the crustaceans.

There is also a need to speed up the test validation process, on the basis of continuing sound science. Additional priorities are:

- a) ethical issues, such as the reduction of the numbers of fish tested via the efficient use of biomarkers, or the replacement of vertebrate test species by the use of mechanistically relevant alternative test species;*
- b) to extend coverage to animal groups of intrinsic ecological and economic importance which currently lack suitable ED tests; and*
- c) getting endpoints/tests accepted for EDCs with specific modes of action for the present guideline species.*

Test strategies, assessment guidelines, and the use of results from OECD test under Community legislation

With respect to testing strategies...

... the most critical question is unresolved: what should be the relevant endpoints and the exact criteria for the final classification of substances as EDs?

said one of the experts. And another one complemented this by the statement:

For regulatory purposes, clear decision criteria, i.e. which positive result(s) in which test(s) are sufficient to decide that a substance is classified as an endocrine disrupter, are required and those are not provided under the OECD EDTA framework. Attempts to reach international consensus on such criteria are probably destined to fail and we assume that it's not in the remit of OECD. There is however a need to carry out work to reach such regulatory decisions at a more regional level, at EU level.

A third expert view contributing to this point was:

There is a desperate need for EU coordination to determine (i) which of the existing OECD ED-screens and test methods should be used within various EU regulatory frameworks for identifying chemicals with ED properties and (ii) which existing OECD test methods should be prioritised for improvement or enhancement with regard to identifying ED-related endpoints, and (iii) which 'new' test methods appear most suitable for use in the EU and which therefore need to be prioritised for international agreement.

Given the time it can take to get fully validated OECD test methods finalised, it may be considered appropriate, in a limited number of cases, for EU regulatory frameworks to specify non OECD validated test methods, although as in cases (ii) and (iii) above – it would be expected that these would be made the subject of OECD work so that they eventually become an agreed and validated OECD test method.

One of the difficulties in this context is the fact that data situations may largely differ between different pieces of EU legislation dealing with EDs:

The current and future development of corresponding OECD guidance documents must properly reflect the fact that there are very different starting points for the assessment of substances under different regulatory regimes. On the one hand, there are data-rich situations, such as for pesticides and biocides, where the assessment in the EU already starts with knowledge from OECD level 4 and 5 assays. On the other hand, there are data-poor situations, such as for many REACH substances, where assessments indeed have to start from level 1 of the OECD framework.

At this point, several experts pointed to the recent proposals for the integration of the output of all the test methods, including the standard supporting and apical regulatory studies that have been made by ECETOC, the German BfR and the UK CoT. They will provide one of the focal points of analyses in the third and final part of this project.

A controversial issue in this context is the question whether the regulatory decision must be based OECD level 5 testing. Some experts argue against this. For example:

- *Validation of in vivo screening tests such as the Uterotrophic Bioassay and the Hershberger Bioassay has taken many years, but up to now, hazard and risk assessment schemes under EU legislation do not make any formal use of them, e.g. as a standard information requirement, probably because REACH was adopted before the final test guidelines were adopted in the OECD. We want to use the information from these assays for hazard classification and preliminary risk assessments, e.g. to enrol substances on the EU*

Community Rolling Action Plan and for the assessment of equivalent concern under REACH Art. 57 (f).

- *To obtain results from definite OECD level 5 in vivo reprotoxicity tests may take several years and may conflict with the EU policy for reducing animal testing. Therefore, the regulatory identification of EDs should not exclusively rely on the results of level 5 testing, but it should make use of all available information from all assays on all levels in a weight of evidence approach. Details of such an approach have been explored in the Nord-Utte EDREG project (Hass et al. 2004)⁷.*

⁷ Hass U, Dalgaard M, Jarfelt K, Kledal TSA (2004) OECD conceptual framework for testing and assessment of endocrine disruptors as a basis for regulation of substances with endocrine disrupting properties. Nordic Council of Ministers, TemaNord 2004:555

2.3 HAZARD BASED GROUPING AND CLASSIFICATION OF EDS

Background

The term *endocrine disrupter* comprises substances and mixtures that act by a variety of mechanisms on an array of endpoints with different potency and specificity.

Questions

- Do different mechanistic groups, such as substances with estrogenic, anti-androgenic, thyroid-disrupting, and other endocrine disrupting activities (“EATO”), require different regulatory approaches?
- Do different endpoints and species-specific types of effects, such as imposex in snails or malformation of reproductive organs in mammals, require different hazard classes for EDS?
- How should different potencies (in terms of dose levels and exposure times at which effects occur) and/or specificities of ED action (in terms of the relation between effect concentrations/doses for ED effects and for other toxic effects) be reflected in hazard classifications for EDS? Should they be reflected at all?

Answers

Almost all experts rejected the ideas of different regulatory approaches for different mechanistic groups and / or different endpoints of EDS. The array of reasons for these opinions is outlined below.

Potency and specificity are generally acknowledged as important criteria for hazard classification. However, the views when and how these criteria should be applied differ widely or are yet unclear as documented by the array of opinions documented below.

Part of the interviewees suggest that EDS should become a hazard class on its own, independent from other categories of the existing hazard classification system for chemicals. Others defeat this idea categorically.

In case of EDS, it may be necessary to consider potency in conjunction with timing of exposure. This was a fourth aspect raised by the experts in response to the interview questions.

Classification by mechanisms and /or endpoints

The array of opinions brought forward against the option of a regulatory classification of EDs by mechanisms and or different endpoints of action reads as follows:

- *Many substances may interfere with more than just a single endocrine pathway. Many estrogens for instance are also potent anti-androgens. Furthermore, many substances may interfere with other pathways than the already well studied EAT types of EDs. In view of these difficulties, substances with different modes of endocrine action should all be lumped together in a single hazard category for regulatory purposes (In relation to REACH Art. 57 (f), pesticides, biocides and also hazard classification. This is a pragmatic way forward.*

Regarding endpoints, the existing evidence does also not yet support the establishment of different hazard classes in relation to EATO. We simply do not have the whole picture that would be necessary for such a differentiation. For instance, when we observe adverse effects in frogs, we currently cannot decide with sufficient certainty whether related effects might also occur in humans or not. Therefore we must take it as a signal for being cautious.

- *Carcinogenic substances are classified as such regardless of the organ or tissue in which they induce tumours. Therefore, drawing a parallel, there is no need to group EDCs by mechanism.*
- *Human Relevance is a key criterion for the assessment of endocrine disrupting properties of a substance for the purpose of human health protection, alongside with some additional considerations on potency.*

Considerations of relevance comprise both the effects observed as well as the mechanisms by which these effects are or might be caused. A priori, assessment procedures for endocrine disrupters should not be restricted to certain mechanisms or endpoints of ED action, nor should such groups of EDs be assessed in different ways. Substances may simultaneously interfere with different parts of the endocrine system and may cause several effects, and furthermore certain mechanisms of action and resulting endpoints may yet be unknown.

Although inappropriate for defining different hazard classes, grouping of EDs by mechanism and/or endpoints is a valuable approach for other purposes, in particular cumulative risk assessments, human relevance assessments of MoAs, and the structuring of assessment work plans.

- *Whether groupings by mechanisms / modes of action (MoA) and/or endpoints of ED action might be useful is questionable. However, the complexity of the matter and the tremendous knowledge gaps in the field speak against such an approach. As an ad hoc statement, information on MoA and endpoints is therefore considered to be important for regulatory decision making, but inappropriate for classification and labelling.*

- *For regulatory action it is not necessary to distinguish between mechanistic groups. Regulatory action is triggered by adverse effects, irrespective of the mechanisms by which they are caused. Nevertheless, mechanistic knowledge is helpful for risk assessments.*

For the regulation it is also not necessary to sub-classify EDs by different specific endpoints. The two examples that you mentioned in your question, imposex in snails and malformation of reproductive organs in mammals, both fall under a general concept of reproductive toxicity. The regulation generally aims to protect from reproductive toxic effects, both human individuals and wildlife populations. However, for the purpose of conducting hazard and risk assessments, such information on species-specific effects is of course very helpful.

- *For regulatory purposes, EDs should be classified as such, in accordance with the IPCS definitions of EDs and adversity. They should not a priori be sub-divided by mechanisms or endpoints, nor should they be considered as a sub-class of reproductive toxicants only.*

Historically, the concern about EDs was initially focussed on substances with EAT modes of action, and with the existing OECD TG 416 as a decisive test for identifying EDs, only substances that are also classifiable as reproductive toxicants can be detected. However, by definition, EDs cover a much wider spectrum of both mechanisms and endpoints.

Apart from reprotox effects, interference of chemicals with the endocrine system may result in a large number of other adverse outcomes, such as effects on the brain (effects on cognitive functions, emotionality, attention, and learning), the immune system, the development of bones and muscle, or the lipid and energy metabolism (which is influenced by brain functions).

Risk assessment must be based on endpoints, as it is not possible to extrapolate from mechanisms to endpoints. However, the TG 416 2-generation assay is clearly not able to cover all the endpoints mentioned, it might even be insufficient for detecting all substances that cause adverse effects by EAT mechanisms. The draft EOGRTS might be an improvement for this purpose but not the ultimate solution for the identification and classification of EDs.

- *I do not see any strong argument for not putting them all together in a single hazard class. The question must be considered in the context of the discussion about thresholds for ED action. Shall EDs be regulated on the basis of the assumption of the existence of a threshold or shall the absence of a threshold be assumed (as for genotoxic carcinogens)? May be, that there are some ED MoAs without a threshold, while for others it is safe to assume a threshold. However, currently we do not know this and hence there is no need to treat them differently.*

The aim of human hazard and risk assessments is to protect every individual, while in ecotoxicology the protection goal is the population or the species. As a consequence, may be that different classes for human and environmental EDs could be a sensible option, but I do not yet have a definite opinion on this point.

- *ED is not a hazard, but a mechanism of action. Therefore grouping can have a role in identifying priorities for further assessment, but not in risk assessment.*
- *Hazard based classification has, up to now, relied on adverse effects and not on the mode of action that lead to the adverse effects. Therefore, in the field of endocrine toxicity, there is no special need for adopting different regulatory approaches according to different “mechanistic groups” (unless there is scientific evidence that one group (eg estrogenic activity) is more hazardous than another group (eg thyroid disrupting activity). However, evidence for such discrimination among different “mechanistic groups” is lacking.*
- *Legislation should always be kept as simple as possible, special roads for different substance classes would make it only more difficult to implement.*

Only a few experts saw relevant counter-arguments in favour of a mechanistic classification, for example:

It is unclear whether endocrine disrupters acting via different mechanisms require different regulatory approaches. Some argue that toxicity affecting receptor signalling differs from other modes of toxicity, eg. action via enzymes, and this is related to the low dose effect issue. Non-monotonic dose-response may be unique to receptor-mediated mechanisms due to feedback loops and this could have significant implications for testing and risk assessment. The low dose issue is puzzling, several well conducted studies have so far failed to observe such low dose responses.

Grouping by mechanism or mode of action could be useful for assessing mixture effects. The assessment of cumulative effects needs to focus on mechanism/mode of action and this should allow us to conduct better risk assessments.

Also only a few experts argued for endpoint related classifications, for example:

Different mechanisms do not require different regulatory approaches,. However different endpoints should be regulated for differently. Different hazard classes would apply depending on whether an endocrine disrupter exerts its adverse effect(s) on human health or the environment.

Views about potency and specificity as classification criteria

The following selection of views about potency and specificity as classification criteria for EDs illustrate the current status of debate on this point.

- *Potency should in general not be reflected in hazard classification, however, for regulatory purposes it can be important to separate “strong” EDs from “weak” EDs in order to prioritise regulatory measures and also seen in the light of combination effects. However, the methodological know-how for such a separation is currently missing and needs to be developed.*

- *Potency is an important criterion for defining different hazard classes of EDs. Provided that the relevance criterion is fulfilled or human relevance cannot be excluded, potency should be used for a distinction between EDs of high concern and EDs of lower concern.*
- *Potency, specificity, and ecological relevance/adversity are important criteria for sound and scientifically justifiable regulatory assessments and differentiated decision making. The need for a differentiated decision making became obvious with the new PPP regulation which established endocrine disruption as a cut-off criterion for the approval of active substances, safeners or synergists. However, some pesticides are actually designed to exert their wanted effects in invertebrates or plants through an endocrine disrupting MoA, such as so-called insect growth regulators, pheromones, and plant growth stimulators. In these cases, a rigorous undifferentiated application of the cut-off criterion may result in questionable decisions or even unfavourable environmental impacts.*
- *Potency should enter classification schemes as an additional criterion. There is the possibility that certain forms of target organ toxicity affect the endocrine system indirectly, and to class substances as endocrine disrupters on that basis is problematic. In principle, endocrine disrupters should elicit endocrine effects as the critical toxicity, i.e. at doses lower than other toxicities.*
- *Potency and specificity in terms of effective exposure levels and their relation to effective doses for other toxic effects are important for hazard classification and regulatory action. We have to think about appropriate settings for such a classification scheme. ECETOC and the German BfR both have presented proposals on the issue. Basically they go into the right direction.*
- *Potency should not enter classification schemes as an additional criterion. There are often bell-shaped dose-response curves – at least their presence cannot be ruled out with the existing testing schemes, and they are suspected to occur with endocrine disrupters – and for that reason endocrine disrupting effects cannot be ignored during classification, even though other toxic effects may occur at lower doses. The problem is that much more testing in the low dose range would be required to rule out with confidence the presence of endocrine disrupting effects, and current testing practice is not equipped to provide the necessary distinctions.*
- *It is probably true that (almost) all substances could produce disruption of the endocrine system at some dose level, just as all substances have toxicity in some form or other under some conditions. So it is crucial that regulatory action is not taken against those substances that disrupt the endocrine system only at 'heroic' doses, or only at dose levels far in excess of those producing other forms of toxicity (on which basis they should be regulated).*
- *Potency is an important aspect, but careful considerations of the following points are necessary:*
 - *Potencies of two chemicals can significantly differ between in vitro, acute in vivo and chronic tests. Relevant for ED classification are only long-term potencies. If long-term*

potencies cannot be determined for practical reasons, potency data from acute tests should be used with great caution.

- Potency indicators (such effective doses, LOAELS or NOAELs) are derived from dose response data, and therefore characteristics of dose-response curves, such as possible deviations from the monotonic type, should be considered. In consideration of present knowledge on EDCs, it is not acceptable to disregard effects that are, e.g., significant “only” for the intermediate dose (1 out of 3-4 doses in TG416, cf. discussions at BfR workshop in Nov 2009).

Specificity, as it is currently discussed, is a too simplistic concept for ED classification. The classical definition of specific effects is that they occur at much lower doses than general toxicity. However, so-called “non-specific effects”, such as reduced weight gain for instance, which usually are considered to indicate “general toxicity”, may be the consequence of an ED effect.

- *Different potencies should not be reflected in hazard classifications for EDs. Hazard classes, such as “C”, “M” and “R”, describe intrinsic properties of chemicals, independent from the dose levels at which these adverse effects occur. Considerations about critical dose levels are an important part of the risk assessment, but not decisive for the hazard classification.*

Different specificities of ED action should be reflected in hazard classifications, if “specificity” is understood in a qualitative mechanistic sense. Subject to expert evaluation, a chemical should not be classified as an ED, if a reproductive effect, for instance, is just a non-specific consequence of a general systemic toxicity and causally not directly linked to an ED MoA.

However; this does not mean that the ED effect must be the so-called “critical effect” or “lead effect” in the meaning used in risk assessment schemes, i.e. no other types of adverse effects are seen below the NOAEL for the critical effect. This would violate the principle that a hazard class, such as “C” or “R”, assigns a specific intrinsic property to a chemical, independent from the question whether one or more other hazardous properties might apply simultaneously.

This last set of statements suggests that EDs should become an independent hazard class. The interviews showed that this is one of the most controversial issues at the current status of debate. Examples are shown in the following.

A separate hazard class for EDs?

Part of the interviewees strongly advocate for an independent hazard classification of EDs, for example:

- *As a regulatory instrument, we want to introduce endocrine disruption as a single new hazard category. An important issue is that we want the category to be sub-divided into suspected and confirmed EDs, but not a priori sub-divided by mechanisms or endpoints.*

In general the classification or non-classification of a substance as an ED should be independent from the classification by other hazard criteria. Thus, specificity in terms of relations between effective doses for different types of hazardous effects is irrelevant for hazard classification (with the comment that endocrine disruption is not as relevant if other severe co-occurring effects will affect animal survival). We do not agree with classification proposals that point into this direction, such as the one presented by ECETOC for instance.

In a mixture, “weak” EDs may still add to the total effect. This is another reason why ED properties should be assessed separately and independent from other hazard criteria.

The opposite standpoint reads as follows:

- *Hazard based effects consider adverse effects independent of their mode of action. Endocrine disruption represents*
 - *neither an independent adverse effect,*
 - *or a new type of toxic property,*
 - *nor a previously undetected hazard.*

Therefore no separate hazard category is required for ED substances. There is no need to have separate regulatory approaches for ED substances as nearly exclusively regulation is based on adversity, not mechanism based. Any change in this paradigm will lead to delay risk management in the area of ED as existing regulation cannot be used and new regulation has to be set in place.

A third part of the interviewees has not yet taken a clear position in this debate:

- *With a mid-term perspective, we are considering whether the establishment of EDs as a new hazard class under the GHS framework for the classification and labelling of substances could be an appropriate step forward. However, the process of forming an opinion on this issue has started only recently. Currently, the potential definition of classification criteria and potential sub-categorizations are totally unclear.*

Potency and timing of exposure

In the context of the debate about potency as a classification criterion, part of the experts pointed to the fact that potency of EDs may critically depend on the timing of exposure. For example:

The definition of EDs requires that the “function of the endocrine system is altered” meaning that not only the extent/dose at which the hormone system is being altered matters, but also that timing matters. It is well known from endocrinology that timing is an important parameter in the

endocrine control of the maintenance of the homeostasis, development and reproduction of organisms. Hence endocrine disruption may be caused by disturbance of the right timing of that control e.g. caused by chemicals inhibiting or activating various receptors, enzymes etc. at especially unfortunate points in time for the organism when such interference triggers abnormal reactions, which may be more or less irreversible. Furthermore, it is also known from endocrinology that dose/concentration response curves may not be monotonic and it is noted that discussion is ongoing in relation to ED and “low dose effects”, inverted U-shaped dose-response curves etc.

Another comment in this context pointed into the same direction:

In terms of potencies, clearly timing of exposure is critical. Critical windows of susceptibility include exposure in utero as well as the pre-pubertal and pubertal periods. As for exposure, negligible exposure as proposed under the PPPR EU regulation for example is unscientific and far too high when compared with endogenous hormone levels or the concentrations of contaminants in water.

2.4 EVIDENCE BASED CATEGORIZATIONS OF EDs

Background

The strength of available evidence for ED activity in humans and/or organisms in the environment may range from indications from computational toxicology such as (Q)SAR, over results from in vitro mode of action screens, results from in vivo studies with laboratory animals, and epidemiological or eco-epidemiological findings, to case observations in humans or wildlife species populations.

Questions

- What should be the guiding criteria for evidence-based categorizations of EDs?
- How should data and knowledge gaps and uncertainties in available data be dealt with?

Answers

The discussion and elaboration of principles and concrete proposals for evidence-based classifications of EDs is still in the beginning. Many interview partners again pointed to the recent proposals made by ECETOC, the German BfR and the UK CoT. In addition to that, most experts outlined some basic ideas or guiding principles, but explained that these still have to be translated into more concrete terms and strategies. It can be expected that some of these will be published in the near future. These will have to be taken into consideration in the third and final part of this project.

As under the previous interview topic, also the question about evidence-based categorisation provoked opposing views about the potential establishment of EDs as an independent hazard class. The views and arguments were the same and are therefore not repeated here.

Views about evidence-based categorisation

The following selection of statements of interviewees illustrates the spectrum of current views about principles of evidence-based classifications of EDs. Currently, lines of agreement and disagreement become not yet very clear on this level.

- *Criteria and procedures for evidence based categorisations of EDs should be applicable to all regulatory classes of chemicals, not just those for which higher tier testing is mandatory, such as pesticides. Therefore we need two basic classes: "Endocrine disruptors" (confirmed) and "Potential endocrine disruptors". The latter will comprise substances with endocrine modulating activity and may be further subdivided into "suspected EDs" and "potential EDs".*

Suspected EDs are those with some evidence of adverse effect due to the endocrine modulating activity. Categorisation of substances as “Potential” EDs provides the necessary stimulus for the generation of more data.

For substances under REACH, data on ED properties do usually not become available automatically. As a consequence, authorities should use in silico methods as a starting point and then continue with a battery of in vitro assays to screen registered chemical substances. In vitro data alone are not appropriate for classification, but are an indication of a need for precaution and can be used for priority setting for further ex vivo and in vivo testing. Burdens for further testing cannot be placed on industry before in silico and in vitro data give strong reasons for concern.

Negative effects on any tier cannot rule out ED properties with safety. Caution is needed when higher tier negative outcomes are used to supersede positive lower tier findings – there may be plausible biological explanations for negative outcomes.

- *For human health endpoints, there is a clear hierarchy of evidence from epidemiological observation of an impact on human health, to experimental evidence in animals in vivo, through to in vitro and in silico approaches.*

For ecotoxicology, however, an adverse effect has to be observed in a population rather than on an individual. There are currently assays that either are validated or in the process of being validated under the OECD framework to address environmental effects of EDCs in fish, birds and amphibians. Beside the need to include molluscs in the battery of tests, these assays will detect effect in individual organisms and not on population, further work is needed to fix criteria in order to prioritize the studies to address the relevance EDCs for a population. It is unclear how this gap can be addressed as evidence of widespread endocrine disruption in wildlife is still wanting.

There are also very different situations in terms of the availability of data depending on the specific use of a substance. Pesticides and Biocides are generally data-rich and it won't necessary be the case under REACH. Therefore an evaluator, working under Pesticides or Biocides, may have evidence of an adverse effect from a one- or two-generation study, and it is then generally possible to find information on the mode-of-action either from industry or in the scientific literature. Under REACH without such data it won't be so easy to detect a possible adverse effect.

From the knowledge of the mechanism of action of triazoles on target organisms for example, it is possible to predict the effect on other environmental receptors. However, their potency may be very different due to the influence of metabolism.

It is also unclear how to balance the relative value of evidence of an adverse effect in an intact organism in one species compared to in vitro evidence of an ED mechanism.

- *The BfR proposal for a hazard-based classification of EDs basically discriminates between evidence from human cases or epidemiological studies and evidence from experimental animal studies. Evidence from animal studies is proposed to be sub-categorised by a combined consideration of exposure duration, endpoint and potency. Evidence from in vitro assays is considered to be insufficient for the classification of EDs, but it may support the assessment of the mechanism of action and human relevance of effects observed in animals.*
- *By definition the classification of a substance as an ED requires both some information on its MoA and some indication on a plausible causal relationship between the MoA and an adverse, ecologically relevant endpoint. Thus, evidences from in vitro and in vivo studies are both required.*

How these abstract criteria should be best translated into more concrete terms and test strategies still need to be elaborated. The work is most advanced for the assessment of endocrine properties of substances in fish. Following the fish example, similar approaches for other ecotoxicological endpoints may be brought forward.

- *In an ideal world, complete knowledge of mechanistic pathways should lead to proper appraisal of endocrine disrupting activities of a chemical, but in the real world, priority reliance has to be on epidemiology, provided that exposure is reliably assessed (generally through proper biomarkers). This obviously applies to compounds to which humans are already exposed. In toxicity testing, two generation studies are key for identification of endocrine disrupters, and for appropriate classification, but there is a real need for more research on those effects (immunotoxicity, epigenetics, metabolic syndrome) not fully covered by current guidelines.*

The outcome of in vitro tests should and can be used to pinpoint mechanisms of endocrine disruption. They can also be used to prioritise chemicals for further in vivo testing, but in vitro test outcomes should normally not be used as a basis for classification or regulation. For the purposes of risk assessment, in vitro assays are not suitable currently, although further developments in this direction are envisaged

- *In general, the identification and assessment of EDs requires a tiered approach such as the Conceptual Framework developed by the OECD EDTA Task Force. However, we should not over-regulate such a process of knowledge generation. For the regulatory decision-making, the crucial point is to define the evidence that is considered to be sufficient for deciding whether a substance is an ED or not. The determinant for reproductive toxicant is the result of an OECD 2-Generation test, in conjunction with some evidence for an endocrine mechanism of action. For non-reproductive pathways, the other endpoints e.g. in 90 day study should be considered in combination with other evidence.*

- *In vitro and in vivo test results should be evaluated together to judge the endocrine disrupting potential of chemicals, and this should be the basis for any categorisations. Positive in vitro test results need to be confirmed in in vivo tests, because of the possibility that due to intervening toxicokinetic processes the chemical in question cannot reach the target tissues. However, in vitro tests provide alerts that need to be ruled out or confirmed in in vivo tests. They are helpful in prioritising chemicals, especially high production volume chemicals for in vivo testing. But in vitro test results alone should not be the basis for regulatory action, e.g. a ban on certain chemicals.*
- *The problem with using evidence from the environment is that although the effects are evident, establishing the relationship between the causal agent and mechanisms (as stated in the definition of an ED) is difficult. Any environmental observations should be supported by laboratory studies where the mechanism and health effects are clearly proven. For example, limb malformations in frogs observed in the USA were attributed to several causal factors (nitrate exposure, UV exposure, parasitism) as well as ED. This highlights the importance of proving the underlying mechanism.*

- *Case by case decisions are generally required; rigid classification schemes are inappropriate.*

Exposure data, both external and internal, as well as kinetic aspects should be taken into consideration. An exclusively hazard based assessment runs into problems. Risk based assessments are preferable.

In vivo effect data from long-term studies should be given more weight than other types of effect data. However, it must be taken into consideration that a negative outcome of a TG 416 assay per se does not rule out any ED potential. Integrated assessments that also consider the potential relevance of ecotoxicological data for human risk assessments as well as potential analogies across mechanistic groups might be advantageous.

In vitro data are insufficient for ED classifications, but provide a tool for priority setting for in vivo testing. Correlations between in vitro and in vivo data appear to be quite good for some anti-androgens, but not for other groups, such as estrogens for instance. More research into mechanistic links is needed here.

- *The highest level of evidence ideally would stem from epidemiological studies, illustrating real world problems and tracking them back to exposure to specific chemicals. This type of study is however hampered by a number of limitations such as latency, for example the difficulty of making an association between fetal exposure and women developing breast cancer in later life. An endocrine disrupter hazard category should be based on the definition, i.e. an endocrine mode of action and adverse effects in vivo. There could be a potential endocrine disrupter category based on a lower level of evidence but there are no plans to do so in the US where there is a concern to avoid alarming people unnecessarily. Nonetheless in*

in vitro testing is great for priority setting. There is a need to investigate metabolism and *in vitro* metabolism bioassays are currently under development.

- Two evidence based categories of EDs should be defined:
 - “definite” or “confirmed” EDs (Cat 1), and
 - “suspected” EDs (Cat 2).

Cat 1, “confirmed EDs”, should be based on good convincing *in vivo* data, and this may include data from the uterotrophic assay with immature rats (OECD level 3)! Subject to case-by-case expert judgement, the general dividing line between convincing and less-convincing *in vivo* data should not be between OECD level 3/4 and 5, but should consider artificial and non-artificial systems and the overall weight of evidence.

Cat 2, “suspected EDs”, should be based on data from *in vivo* screening assays in combination with good supportive evidence from other studies, such as toxicokinetic information, *in vitro* and *in silico* findings and read-across from similar substances.

- Screening tests can only explain mechanisms. Per definition, ED substances cause effects in intact organisms. The organism is complex and in many instances can adapt to environment. It is therefore not sufficient to have *in-vitro* screens only.

Knowledge gaps and uncertainties can only be closed with additional data. Results from *in vitro* screening tests are not sufficient to conclude possible ED effects. In specific cases, especially if no higher Tier data are available, they can either guide priority setting for further testing or give additional information bringing clarity to some observed effects in higher tier testing. Thus they support *in vivo* studies and help to explain mechanisms behind observed effects. *In-vitro* studies cannot substitute *in-vivo* studies, even if there is a lack of data. (definition: effect seen in intact organisms)

The guiding criteria for evidence-based categorizations of EDs should be based, at least, on the integration of *in vitro/in vivo* mode of action data supported by data showing adverse effects in supporting/apical *in vivo* studies. These guiding criteria have been well described in the ECETOC Technical Report no. 106 “Guidance on identifying endocrine disrupting effects” (2009) and in a recent publication (Bars R. et al., 2010. Regul. Toxicol. Pharmacol. In press).

2.5 CUMULATIVE EXPOSURE AND MIXTURE EFFECTS OF EDS

Background

Humans and organisms in the environment may be exposed to multiple EDs from different sources via different routes, simultaneously or sequentially. Some EDs may jointly act by a common mechanism or they may contribute to a common endpoint by different mechanisms. The Council has called upon the Commission *to make recommendations as to how exposure to multiple endocrine disruptors should be further addressed within relevant existing Community legislation* (Council of the European Union, Document 17820/09, Annex, point 7).

Questions

to EU MS agencies:

- What could be sensible options for improved protection of humans and the environment from potential combination effects of EDs within the framework of existing Community legislation?
- Do the existing pieces of EU chemicals legislation provide a sufficient basis for adequate protection from hazards and risks posed by ED mixtures or are new complementary approaches required? If new approaches are required, what elements should they comprise?

to other organisations:

- Are there activities in your authorities to consider cumulative exposures and mixture effects of EDs, and if yes, which?
- Do the existing pieces of legislation in your country provide a sufficient basis for adequate protection from hazards and risks posed by ED mixtures or are new complementary approaches required? If new approaches are required, what elements should they comprise?

Answers

Almost all interview partners stated that they consider the assessment of cumulative exposure and mixture effects as a topic of high priority. However, they also almost uniformly stated that this is a general topic and not specific for ED assessment. Consequently the range of views and opinions that came out from the interviews basically reflects the state of discussion that has been documented in the State-of-the-Art Report on Mixture Toxicity and the sub-subsequent discussions of that report. In

fact none of the statements brought up any new arguments, views, or activities on that point, and in particular none of the statements highlighted any aspect that is specific for EDs. As consequence, no detailed documentation of the spectrum of views obtained is included in this overview.

2.6 REGULATORY ASSESSMENT OF EDs – WHAT SIGNIFIES AN EQUIVALENT LEVEL CONCERN?

Background

Under the European REACH legislation, substances become subject to authorization if they are classified as CMRs cat 1 or 2, PBTs, or vPvBs according to criteria specified in the legislation. EDs (and other substances) will also be subject to authorization if *there is scientific evidence of probable serious effects to human health or the environment which give rise to an **equivalent level of concern** to those classified as CMRs cat 1 or 2, PBTs, or vPvBs* (Regulation (EC) No 1907/2006, Article 57).

Under the new European Regulation on plant protection products, the use of a chemical as an *active substance, safener or synergist* in a plant protection product (PPP) *shall only be approved if, (...), it is not considered to have endocrine disrupting properties that may cause adverse effect in humans* (Regulation (EC) No 1107/2009, Annex II, 3.6.5).

Questions

to EU MS agencies:

- What is an interpretation of the term **equivalent concern** that is consensually acceptable as an adequate reflection of the legislative intent and which provides precise clear-cut decision criteria that are consistently applicable across different pieces of EU legislation?

to other organisations:

- Does your authority/organisation regard EDs as presenting concerns equivalent to carcinogens, mutagens, reproductive toxicants or persistent and bioaccumulating agents?

Answers

Taken together, the answers obtained from interview partners provide a list of arguments and concerns why EDs may exert severe, non-reversible effects on humans or the ecosystem. However, currently it remains largely unclear where the exact borderline between high and low concern should be drawn. Expert judgement and case-by-case decisions remain inevitable.

Views and arguments about criteria for “equivalent level of concern”

The following selection of statements characterises the spectrum of views and opinions about “equivalent level of concern” that came out from the interviews:

- *equivalent level of concern*
 - *Serious effects, irreversible effects, and unknown patterns of effects that are difficult to discover*
 - *Every exposure is potentially problematic*
 - *Survival of the organism/population – effects on next generation*
 - *Timing of exposure may be critical (time windows during development).*
 - *There may be a very long latency between exposure and adverse health outcomes.*
 - *Even by effective intervention, results may first be seen in about 20-30 years.*
 - *In relation to the issue simultaneous exposure: dependant on ADME of the substances in question, what in practice should be regarded as simultaneous.*
 -
- *From a human health perspective, the equivalent level of concern requires that approaches to develop criteria for classification and categorisation should be consistent with those applied for CMR with Cat1a, Cat1b and Cat2 endocrine disrupters. The same parallel can be drawn for environmental EDs and the need for approaches consistent classification as PBTs and vPvBs where there are no sub-categories, Deriving PBT-like criteria will be particularly challenging because an effect at the population or ecological community level depends not only on the severity of the effect and the number of species affected but also on the timing of exposure, pseudo-persistence of compounds or repeated rather continued exposure. There is a need to develop and set specific environmental criteria which may trigger ED classification, these criteria should take into account the effect at population level and the potency of the substance.*
- *The assessment of an equivalent level concern, as compared to CMRs cat 1 or 2, PBTs, or vPvBs, must basically refer to the hazardous potential of a substance. The details, however, are currently subject to different interpretations. It is unclear where the exact borderline between high concern and lower concern should be drawn, since it is unclear whether the assessment should include a reflection of the exposure situation and thereby refer to an equivalent risk and not only to equivalent hazardous properties.*

The assumption that threshold values cannot be derived is generally applicable to genotoxic carcinogens and mutagens, but it is not generally true for reproductive toxicants and endocrine disrupters. Consequently, the assessment of an equivalent level concern cannot be based on threshold considerations.

- *Exposure to EDs poses risks that warrant an equivalent level of concern to CMR, PBT and vPvB. It is important to bear in mind that ED may elicit carcinogenic, epigenetic and reproductive toxic effects via modes of action that are different from those operating for “classical” carcinogens or reproductive toxicants, but that justify an equivalent level of concern.*
- *Apart from PBTs and vPvBs, the precise and clear-cut criterion for identifying a substance of very high concern, which becomes subject to authorisation under REACH, is the positive outcome of a 2-generation study that leads to the classification as a CMR cat 1 or 2. EDs are just one out of several groups of substances that are covered by this criterion. Hence, the additional term “equivalent level of concern” does not provide any obvious added value for the regulatory decision-making for the reproductive toxicants. It might be necessary to include other ED pathways that raise concern and can be defined for the regulatory purpose. Thus, the legislative term “equivalent level of concern” is something that leaves space for future developments. It provides a mandate for appropriate regulatory action in case of new, currently unforeseeable insights or events.*
- *Exposure to EDCs poses risks that warrant an equivalent level of concern to CMR, PBT and vPvB, mainly because many EDCs are also CMRs. However, what precisely should constitute an “equivalent level of concern” should not be defined to restrictively, and should be judged on a case-by-case basis. Potential for endocrine disruption, e.g. with chemicals currently placed in category 3 or non-classified substances with endocrine properties in combination with P or B, can also justify an equivalent level of concern.*
- *“Equivalent concern” should reflect a hazardous property that, if exposure is not adequately controlled, could produce a risk of effects as unacceptable to society as cancer, genetic damage or adverse reproductive effects. Perhaps a broader debate is needed on what such effects might be?*

In terms of ecotoxicology, adverse effects on population recruitment or stability in a statistically significant proportion of the study population, with regard to reproduction, development and disease incidence. Potency could also be a very important factor here.
- *EDs represent an “equivalent level of concern” as CMRs. There is no official justification of an “equivalent level of concern”, but the following arguments should be considered:*
 - *Natural hormones are active at very low levels. EDs might only be active at levels that are an order of magnitude higher. However, in comparison to the actual exposure this still might be low levels that give reasons for concern.*

- *Hormones play a crucial role during development. Interference of EDs with the hormonal system may result in persistent and irreversible changes of structure and function.*
 - *Hormones act on a multitude of different tissues. Interference with the hormone system may thus affect a multitude of tissues, and not just a single organ.*
-
- *equivalent level of concern*
-
- *Endocrine disruption is a mode of action of toxicity where adverse effects are manifested in reproductive and/or subchronic/chronic/carcinogenicity studies. These adverse effects are already considered by existing classification scheme so having a separate concern for EDs is redundant. In other words reproductive toxicity should not be considered more critical when it is mediated by an endocrine mechanism than when it is mediated by a non-endocrine mechanism of toxicity.*
-
- *EDs may give rise to an equivalent level of concern as PBT chemicals. As with PB(T) chemicals, we feel that the justification of equivalent concern is based on the uncertainty of the risk assessment and the potential for ongoing irreversible effects if controls over the chemical are found to be inadequate. For example, animal experiments show that effects due to endocrine disruption may only become apparent at maturity, many years after the harmful in-utero exposure took place. Therefore, like PB(T) chemicals, if EDCs have been inadequately regulated, adverse effects may be seen for many years to come, and future generations may be adversely affected.*

2.7 PRAGMATIC DEFAULT ASSESSMENT CRITERIA FOR EDS

Background

The new PPP legislation requires the development of *specific scientific criteria for the determination of endocrine disrupting properties* until December 2013 and simultaneously sets out interim criteria ((Regulation (EC) No 1107/2009, Annex II, 3.6.5):

Pending the adoption of these criteria, substances that are or have to be classified, in accordance with the provisions of Regulation (EC) No 1272/2008, as carcinogenic category 2 and toxic for reproduction category 2, shall be considered to have endocrine disrupting properties.

In addition, substances such as those that are or have to be classified, in accordance with the provisions of Regulation (EC) No 1272/2008, as toxic for reproduction category 2 and which have toxic effects on the endocrine organs, may be considered to have such endocrine disrupting properties.

Question

to EU MS agencies:

- Would it be sensible to adopt these interim criteria as permanent ones?

Answers

One of the experts defended the interim criteria as pragmatic and potentially also permanently applicable approach. The others rejected them for a number of reasons that are compiled below.

Arguments

The expert arguing in favour of the interim criteria stated:

The interim criteria for hazard-based cut-offs for endocrine disruption as defined in the new PPPR could be made permanent – this would provide a good and pragmatic solution. However, it has to be borne in mind that not all CMR effects are endocrine mediated.

The others provided the following counter-arguments

- *Adoption of the interim criteria as permanent ones would be nonsense for the following reasons:*

- *Potent alkylating substances will also cause cancer and reprotox without being endocrine disrupters.*
- *Several endocrine disrupting effects will not be covered by these criteria.*
- *We would not recommend the interim default assessment criteria, being made permanent. While it is true that some EDs will be both Cat2 reproductive toxicants and Cat2 carcinogens, there are also compounds that are C2R2 whose mode-of-action is entirely different from endocrine disruption.*
- *An adoption of the interim criteria, which have been laid down in the new PPP regulation for the assessment of EDs, is not acceptable. This was a consensual outcome of the scientific expert workshop on the issue held by the BfR in November 2009.*

Furthermore, there are strong reasons to assume that the application of the interim criteria might result in false assessments of endocrine disrupting properties of substances, since modes/mechanisms of action are not considered.

False positive assessments of endocrine disrupting properties of substances resulting from the application of the interim criteria under the new PPP regulation may provoke court cases. Therefore, there is an urgent need to develop science-based criteria as soon as possible.

- *The interim criteria for hazard-based cut-offs for endocrine disruption as defined in the new PPPR should not be made permanent, because endocrine disrupters operate by mechanisms very different from those relevant to CMR. For example, numerous carcinogens are not endocrine disrupters. In classing chemicals as endocrine disrupters, endocrine effects that are secondary to the main effects of the chemical in question should be excluded, and the demand that endocrine effects should occur as the critical toxicity is crucial. In this respect, the scheme proposed by BfR provides a viable approach.*
- *It would absolutely not be sensible to adopt the interim criteria as permanent ones*

Endocrine disruption is a mode of toxic action. CMR categorisation is based on weight of evidence for human relevance of a toxic effect observed (usually in animal experimentation). For example, a reactive chemical might well produce genotoxicity, leading to cancer, and damage reproductive organs/processes, such that cat 2 C and R status is merited. This would have nothing to do with ED. Hexavalent chromium would be one such example.

The above interim criteria will include secondary or indirect effects on the endocrine system; such as increased clearance of steroid hormones via the induction of detoxification enzymes and/or apoptosis on cells of the endocrine system. It's important to define the difference between a chemical that disrupts endocrine function (by altering production, transportation and binding) and one that acts via toxic or secondary mechanisms.

- *A combination of "C" Cat 2 and "R" Cat 2 does not make an ED. Some genotoxic compounds are both "C" and "R", but it would be non-sense to conclude that they are an ED*

2.8 COMPLEMENTARITIES OF NATIONAL APPROACHES AND COMMUNITY APPROACHES

Background

In addition to the *Community strategy for endocrine disruptors*, several EU MS have established national initiatives, programs, or strategies for research and/or regulatory activities in the ED field.

Questions

to EU MS agencies:

- What are the specific aims and approaches of recent and forthcoming activities on the MS level? What are significant results from past activities? What are or what may be potential consequences for complementary action on the Community level?

Answers

All interview partners kindly gave an overview over their past activities and provided corresponding published material. The resulting picture is basically identical with the overview that has been prepared in parallel to this project by the EFSA Endocrine Active Substances Task Force and that has been published in November 2010⁸. As a consequence, this is not repeated in this overview here.

Many interview partners kindly also provided some insights on their current and planned activities in the field. As a result, significant new material can be expected to appear in the second quarter of 2011 and this will have an impact on the work in the third phase of this project. Further details cannot be given in this overview without violating the confidentiality that was assured to all interviewees.

⁸ EFSA (European Food Safety Authority) (2010) Scientific report of the Endocrine Active Substances Task Force. EFSA Journal 2010;8(11):1932