Scientific Colloquium Series of the European Food Safety Authority N° 17 - June 2012

SUMMARY REPORT[1]

EFSA SCIENTIFIC COLLOQUIUM XVII

Low-dose-response in toxicology and risk assessment

14 – 15 June 2012
Parma, Italy

[1] Question No EFSA-Q-2012-00759.

Suggested citation: European Food Safety Authority; EFSA Scientific Colloquium XVII: Low-dose-response in toxicology and risk assessment. Supporting Publications 2012:EN-353. [27 pp.]. Available online: www.efsa.europa.eu/publications

Acknowledgements

EFSA thanks Anna F. Castoldi, Jean-Lou Dorne, Djien Liem, Robert Luttik, Luc Mohimont, Iona Pratt and Stef Bronzwaer for their preparatory work as the Organising Committee, as well as Alexandre Feigenbaum and Robert Luttik for acting as overall chair of the Colloquium. EFSA would also like to thank Wim Mennes for his role as overall rapporteur and for providing the take-home messages at the end of the Colloquium; and Claudia Heppner, Robert Luttik, Maricel Maffini, David Bell, Jason Aungst, Dieter Schrenk, Laura Vandenberg, Iona Pratt, Edward Calabrese for their presentations at the opening session; Trine Husøy, Ursula Gundert-Remy, Christophe Rousselle, Fernando Aguilar for having acted as rapporteur of the discussion groups, and Susanne Hougaard Bennekou, George Loizou, Paul Brantom, Anthony Hardy for chairing the discussion groups.

Table of contents

1. INTRODUCTION

"*All substances are poisons. It's the dose that makes the poison*". This famous statement by Paracelsus (1493-1541) is the basis for a fundamental concept in toxicology and risk assessment: the individual response of an organism to a chemical increases proportionally to the exposure (dose). Also, it is generally accepted that for most chemicals there is a threshold dose below which there is no adverse effect.

In recent years, the classical (monotonic) dose-response paradigm has been challenged by the so-called 'low dose hypothesis', particularly in the case of endocrine active substances. According to this hypothesis, a number of chemicals, in particular hormonally active agents, often also referred to as endocrine disruptors or endocrine active substances[2] may exert "low dose effects", i.e. in the range of typical human exposure, which are not present at higher doses, and which may display a non-monotonic dose-response (NMDR) profile, e.g. U-shaped, inverted U-shaped. According to the NMDR hypothesis, a non-monotonic relationship between dose and effect would not allow, for a given effect, a simple monotonic extrapolation from high to low doses during risk assessment of those substances.

"Low-dose effects" have been defined as any biological change occurring in the range of typical human exposures or at doses below those typically used in the standard testing protocols. Some chemicals with hormone-like activity, e.g. some pesticides, dioxins, polychlorinated biphenyls (PCBs), polychlorinated dibenzofurans (PCDFs) and bisphenol A (BPA), have been claimed to produce low-dose effects. Although some effects reported at low doses have been suggested to show non-monotonic dose response curves (NMDRC), non-monotonicity is not synonymous with "low-dose effects". In a NMDRC, the slope of the curve changes sign somewhere within the range of doses examined. Therefore, in those conditions a safe dose level determined from high dose toxicity testing would not guarantee safety at lower untested doses that may be closer to current human exposure levels.

As yet no scientific consensus has been reached as to the validity of the studies supporting the low dose hypothesis. However, a number of new studies have been published that may provide further support for this hypothesis. It follows that there is high scientific and public interest on how the low dose hypothesis can be taken into account when assessing chemical risk and food safety.

It should be noted that a detailed evaluation of the scientific evidence supporting or refuting the validity of these two hypotheses as well as the discussion of particular case substances (for example BPA) were considered to be outside the scope of this colloquium.

The objective of the Colloquium was to bring together international experts from different sectors for a scientific debate on the current state of the art in low dose-response in toxicology and to identify ways of further enhancing the process of food and feed risk assessment in the European Union (EU). Over two days, 100 scientific experts exchanged views and debated the possible health effects of low levels of certain chemicals and the current and future challenges these pose for risk assessment. The 17th Scientific Colloquium organised by the European Food Safety Authority (EFSA) attracted risk assessors, risk managers, scientists and stakeholders from 21 countries, including 12 EU Member States, 4 EU Candidate Countries, Japan, Norway, Russia, Switzerland and the United States. The Colloquium, chaired by Robert Luttik and Alexandre Feigenbaum, welcomed toxicologists,

[2] The terms endocrine disruptors, endocrine disrupting chemicals, and endocrine active substances are often used interchangeably by different professional groups and different geographical locations, although not carrying the exact same meaning. It was not in the objectives of the Colloquium to agree on the use of one specific definition, and this Summary Report reflects the different terms used by different speakers and rapporteurs.

endocrinologists and biochemists from academia, industry and public health authorities, including representatives of several European national competent authorities, the European Commission (EC), the Joint Research Centre (JRC), the Scientific Committee on Emerging and Newly Identified Health Risks (SCENIHR), the European Chemicals Agency (ECHA) and the U.S. Food and Drug Administration (FDA).

The opening plenary session was dedicated to key-note lectures (see abstracts below). These lectures briefed the Colloquium participants on the current debate and provided a good background for contributions to the discussion groups. The presentations included:

- Welcome and introduction to EFSA (Claudia Heppner)
- Objectives of the Colloquium (Robert Luttik)
- Report on Pew, Nature and IFT cosponsored workshop on Non-Monotonic Dose-responses: Relevance and Implications for Food (Maricel Maffini)
- Nature of an effect: adverse or non-adverse? (David Bell)
- Dose-response relationships: biological and modelling aspects (Jason Aungst)
- Low dose effects: is the lowest the most relevant? (Dieter Schrenk)
- When the dose doesn't make the poison: low dose effects and endocrine disrupting chemicals (Laura Vandenberg)
- Low dose effects - impact for risk assessment (Iona Pratt)
- The Hormetic Dose-response (Edward Calabrese, via video-conference).

Following this introductory session participants divided into four discussion groups, each focusing on a specific key issue: the nature of an effect and the assessment of adversity; dose-response relationships; the evidence for NMDR curves; and the challenges for risk assessment.

2. ABSTRACTS OF SPEAKERS IN OPENING PLENARY SESSION

Report on Pew, Nature and IFT cosponsored workshop on Non-Monotonic Dose-responses: Relevance and Implications for Food

Maricel Maffini, *The Pew health Group, United States*

In 2010, the Pew Health Group launched its Food Additive Project. Its purposes are to: (1) conduct a comprehensive analysis of the existing regulatory program; (2) determine whether that system ensures that chemicals added to food are safe as required by law; and (3) develop policy recommendations. Through a transparent process that engages industry, academic, government, and public interest stakeholders, project staff consult with a team of expert advisors, hold workshops, and publish peer-reviewed journal articles. More information on the initiative available from www.pewhealth.org

The project has convened a series of meetings of scientists that focused on the hazard identification and characterization of chemicals added to human food, dietary exposure assessment and potential policy solutions to issues identified throughout our assessment. As a result of the discussions on hazard assessment, it became clear that there was disagreement over the relevance of the shape of dose–response curves and low dose effects, and that these issues deserved more discussion. On April 2012, Pew convened a meeting titled "Non-monotonic dose-responses: Relevance and implications for food" attended by scientists from academia, regulatory agencies, public interest groups and risk assessment community. The goal was to start a dialogue about the relevance of scientific evidence on endocrine disruption, not to reach consensus. Participants acknowledged that this dialogue was long overdue.

It is apparent to us that (1) the potential public health implications of non-monotonicity at doses relevant to human exposure are significant enough to warrant making the issue a priority; (2) there is a need to improve the interdisciplinary communication of endocrinologists, toxicologists, and risk assessors to better evaluate these implications; and (3) addressing non-monotonicity will likely require a rethinking of most current risk assessment approaches.

Nature of an effect: adverse or non-adverse?

David Bell, *European Chemicals Agency, Finland*

When addressing 'effects' seen in toxicology studies used for regulatory purposes, it is first important to consider the experimental context. The quality of scientific studies should be assured, with, inter alia, validated experimental procedures and with reported results accurately and completely reflecting raw data. Further the experimental design should be robust, with reliable methodology, an understood output, and, where appropriate, use of multiple exposure levels to enable evaluation of the relationship between dose and effect. Thus in REACH (Regulation (EC) No 1907/2006), toxicological studies must be performed in compliance with the principles of Good Laboratory Practice (Article 14(4)) and in accordance with specified test methods (Article 13(3)), although Annex XI provides specific conditions for the use of existing data which is not in accordance with Good Laboratory Practice or a recognised test method. When evaluating an 'effect', it is first important to determine whether it is substance-related or not. The WHO International Programme on Chemical Safety has published a definition of adversity. However, it remains necessary to consider a variety of issues, amongst which is the biological plausibility, in considering whether an effect is adverse or not.

Dose-response relationships: biological and modeling aspects
Jason Aungst, *U.S. Food and Drug Administration, United States*

Dose-response is the relationship of an effect due to a chemical or other compound over a range of dose levels. The dose-response relationship is important in toxicology to build an understanding of the integrated biological processes underlying a response. The progression of and reproducibility of an effect over multiple doses can allow extrapolation of the potential for, or lack of, effects at other doses. In this manner, a proper dose-response analysis can contribute to endpoint validation and hazard identification and is an essential component of a regulatory safety assessment. Additional methods are available (e.g. pharmacokinetics, PBPK models) to enhance interpretation of dose related effects and decrease uncertainty in extrapolation from a dose-response relationship when characterising risk. Biological variability and analytical uncertainty are inherent in interpretation of a dose-response relationship. Examination of multiple endpoints in a dose-response assessment, comparison of dose-response data across chemical class, and methods to better characterize dose can significantly reduce this uncertainty, identify data gaps, prioritize testing, and predict the potential for additional effects at other doses.

Low Dose Effects: Is the Lowest the Most Relevant?
Dieter Schrenk, *University of Kaiserslautern, Germany*

The lowest dose causing adverse effects is not necessarily the most relevant. Understanding of the Mode of Action (MoA) of a chemical is the important requirement for any decision on the relevance of a given effect/endpoint for the selected target (e.g. humans). In order to achieve an understanding of the MoA, a sufficient amount and/or quality of scientific studies including mechanism-targeted studies, is required. In any risk assessment, the type and quality of the literature eventually considered should be defined, if possible, in advance. After selection of studies based on the aforementioned quality criteria, the most sensitive endpoints are selected and scrutinised according to the Hill criteria (Hill, 1965). A mode-of-action-analysis is aimed at identifying key events, associated events, and modulating factors in experimental models (rodents, etc.) according to the IPCS framework. A targeted (human) relevance decision can be made based on a decision tree as previously suggested by Boobis et al. (2008). Dose-response considerations (comparison of model vs. target) for the critical MoA are made using mathematical models. In case, target (human) relevance is accepted, dose-response considerations made modify the risk assessment, i.e. the numerical outcome of a risk descriptor. The aforementioned stepwise procedure is illustrated using the example of carcinogenicity of dioxins. It is suggested that there is no reason to believe that the MoA of dioxin causing liver tumors in rodents is fundamentally different from a MoA in humans, while dose-response differences between rodents and humans may modify human risk assessment.

When the dose doesn't make the poison: low dose effects and endocrine disrupting chemicals
Laura Vandenberg, *Tufts University, United States*

For decades, studies of endocrine disrupting chemicals (EDCs) have challenged traditional concepts in toxicology, in particular the dogma of "the dose makes the poison", because EDCs can have effects at low doses that are not predicted by effects at higher doses. In our recent review (Vandenberg et al., 2012), we discussed in detail two major concepts in EDC

studies: "low dose" and non-monotonicity. In 2001, "low dose effects" were defined by the US National Toxicology Program as those that occur in the range of human exposures, or effects observed at doses below those used for traditional toxicological studies. We reviewed the mechanistic data for low dose effects and used a weight-of-evidence approach to analyze five examples from the EDC literature. I will discuss two of these examples, the effects of atrazine on sexual differentiation in amphibians, and the effects of BPA on the mammary gland in rodents. Additionally, we explored non-monotonic dose-response curves (NMDRC), defined as a non-linear relationship between dose and effect where the slope of the curve changes sign somewhere within the range of doses examined. We provided a detailed discussion of the mechanisms responsible for generating these phenomena, plus hundreds of examples from the cell culture, animal and epidemiology literature. We have illustrated that non-monotonic responses and low dose effects are remarkably common in studies of natural hormones and EDCs. Whether low doses of EDCs influence certain human disorders is no longer conjecture, as epidemiological studies show that environmental exposures to EDCs are associated with human diseases and disabilities. Our review of over 840 references concludes that when NMDRCs occur, the effects of low doses cannot be predicted by the effects observed at high doses. Thus, we have proposed that fundamental changes in chemical testing and safety determination are needed to protect human health.

Low dose effects - impact for risk assessment

Iona Pratt, *Food Safety Authority of Ireland, Ireland*

Risk assessment of chemicals in food is based on the paradigm of hazard identification, hazard characterisation, exposure assessment and risk characterisation. Hazard characterisation involves evaluation of the relationship between the level of exposure and an adverse response in standardised animal toxicological studies. For thresholded effects, the No-Observed-Adverse-Effect-Level (NOAEL) or the Benchmark Dose (BMD) in the study can be used to derive (by application of an uncertainty factor) a health-based guidance value (e.g. ADI or TDI). The ADI / TDI represents an exposure level at which it can be concluded with reasonable certainty that no adverse effects will occur in a human population exposed to the chemical for their lifetime. In the case of a NMDR curve the traditional NOAEL / BMDL point of departure arguably cannot be used to derive a health-based guidance value. This reflects the uncertainties regarding identification of an exposure level at which it can be concluded with reasonable certainty that the risk for the exposed population is minimal / non-existent. An additional issue is the possibility that there may be critical windows of exposure for the induction of adverse health effects. It may not therefore be possible to identify a health-based guidance value that is appropriate for the lifetime of the entire population. These considerations could dictate a need for new risk assessment approaches or modifications of existing approaches. Possibilities include the use of additional uncertainty factors, application of the Margin of Exposure (MoE) approach used for the risk assessment of (non-thresholded) genotoxic carcinogens, low dose extrapolation. Consideration of the impact of low dose-responses on the risk assessment process will require careful evaluation of the shape of the dose-response curve, scientifically-based decisions regarding the adverse nature of the effects seen, and consideration of study designs incorporating endpoints beyond current OECD methods.

The Hormetic Dose-response

Edward J. Calabrese, *University of Massachusetts, United States*

This presentation provides an assessment of hormesis, a dose-response concept that is characterised by a low-dose stimulation and a high-dose inhibition. It will trace the historical foundations of hormesis, its quantitative features and mechanistic foundations, and its risk assessment implications. It will be argued that the hormetic dose-response is the most fundamental dose-response, significantly outcompeting other leading dose-response models in large-scale, head-to-head evaluations used by regulatory agencies such as the EPA and FDA. The hormetic dose-response is highly generalisable, being independent of biological model, endpoint measured, chemical class, physical agent (e.g. radiation) and interindividual variability. Hormesis also provides a framework for the study and assessment of chemical mixtures, incorporating the concept of additivity and synergism. Because the hormetic biphasic dose-response represents a general pattern of biological responsiveness, it is expected that it will become progressively more significant within toxicological evaluation and risk assessment practices as well as having numerous biomedical applications.

3. SUMMARY OF DISCUSSION GROUP OUTCOMES

Following the introductory presentations, participants split into discussion groups to debate specific issues in more detail. Participants were provided with guidance on the remit of the discussion groups via a presentation by Stef Bronzwaer. Before the Colloquium all participants had received briefing notes, including selected references for further background, so as to be prepared for an interactive exchange of views and expertise during the Colloquium. Participants were divided, based on their preferences, into four groups to allow parallel discussion groups. A summary is presented below. These summaries are structured following the short set of discussion points that had been formulated in the briefing notes for each discussion group.

3.1. Discussion Group 1 - Nature of an effect: Adverse or non-adverse?

Chair: Susanne Hougaard Bennekou – Rapporteur: Trine Husøy

3.1.1. What experimental evidence would be necessary to define adversity for low dose effects and non-monotonic dose-responses?

The group participants considered that current toxicological testing protocols can be used to study the effects of chemicals with low dose effects and/or NMDRC. The main change in testing strategy is that we need to test more doses in order to identify such effects, especially in the low dose area. If the magnitude of the anticipated effect is small, the statistical power of the study should be considered. For endocrine disruptors, which are reported to have effects at low doses, the group concluded that all organs should be considered as target organs for adverse effects, as hormones may affect all tissues.

Results from epidemiological studies are also useful to identify adverse effects at low doses, but the causality is difficult to determine. To select the proper low dose levels for a specific chemical, human exposure should be taken into account. *In vitro* studies can be used for priority setting and to study the mode of action. They can also be useful to better identify and characterize adverse effects. The group considered that in order to provide definitive proof of adversity, the experimental results need to be reproducible.

3.1.2. Is the working definition of adversity for low dose effects, together with the factors to be considered, still valid?

Lewis et al. (2002) defined an adverse effect as "a biochemical, morphological or physiological change (in response to a stimulus, in this case the chemical substance) that either singly or in combination adversely affects the performance of the whole organism (the test species) or reduces the organism's ability to respond to an additional environmental challenge". The group considered that this definition of adversity does not need to be different for low dose effects.

At low dose levels the biological response signal may be very weak and difficult to detect reliably. To decide whether an effect is adverse or not can be problematic for effects of small magnitudes. The group emphasised that it is important to consider biological plausibility.

The group considered that it is unclear whether significant adverse effects are missed when high dose levels only are tested. It is also unclear whether low dose effects are different from those observed at high doses.

3.1.3. Would the NOAEL / BMDL concept for defining a non-adverse PoD still be applicable for low dose non-monotonic dose-response effects?

The benchmark dose lower bound (BMDL) can be used to define point of departure (POD) for non-monotonic dose-responses if we have reliable data that sufficiently describe the curve and the uncertainty. The different parts of the curve can be analysed separately. An adequate BMD response (1%, 5%, 10%) has to be decided case by case. When studies with only a few data points would be the critical ones for the risk assessment, the BMD approach cannot be used to define a PoD for non-monotonic dose-responses.

The group considered that it may be difficult to identify a NOAEL from a NMDRC. Dependent on the dose-response curve, several NOAELs may exist. Human exposure data may be used to indicate which NOAEL to use. However, since the whole dose-response curve very often is not known, there will be large uncertainties in defining a NOAEL for NMDRC.

3.1.4. Defining data gaps to be filled in order to establish a point of departure that can be used in the risk assessment of low dose non-monotonic dose-response effects

The group considered that to establish a PoD for NMDRC well described dose-response curves are needed, with more doses tested in the low dose range. The low doses tested should be decided from human exposure and mode of action of the chemical. There are no standard requirements for end points tested, and the end points tested have to be decided case by case dependent on the mode of action. *In vitro* models can be used to produce hypotheses on mode of action for new chemicals. Development of validated *in vitro* models is needed. Studies on toxicokinetics including internal dose measurements have to be investigated, and would help in producing hypotheses on mode of action (MoA). PBPK modelling could be considered. The group also challenged the need to test at high doses which are not relevant to any human exposure scenarios and therefore provide data that is not relevant to protecting human health.

3.1.5. What are the implications of using non validated experimental animal models in defining adversity for low dose, non-monotonic dose-response effects?

In many areas of research on low dose, NMDR effects, validated studies are not available and the assessment may need to rely on non-validated studies based on good science. Experiments from new non-validated models can be used when they are properly described, variability is understood and controlled, and when reproducibility and reliability is established. For non-validated studies it is necessary to demonstrate that the results can be repeated. The model should be based on good science and should be demonstrated to be relevant. Poorly described experiments in non-validated models should not be used. One-dose level studies at low dose are of limited use, and should be repeated with several doses. However, they can be used to inform hypotheses on mode of action.

3.2. Discussion Group 2 - Dose-response relationships

Chair: George Loizou - Rapporteur: Ursula Gundert-Remy

3.2.1. A discussion of the toxicokinetic and toxicodynamic aspects of dose-response in biology and toxicology

The profile of a dose-response curve is determined by both the toxicokinetics and the toxicodynamics of the substance under evaluation. Hence, non-linearities in the toxicokinetics which might be dose dependent certainly influence the form of the dose-response relationship.

Current modelling of dose-response curves is data driven. This means that a mathematical function is fitted to describe the empirical data. The function is therefore a mathematical expression rather than a description of the underlying biological and pathophysiological mechanisms. Extrapolation to doses outside the range described by empirical data is determined by the mathematical function which has been used to describe the data and has a high uncertainty because it is not based on a validated biological model. On the other hand, an observed non-monotonicity without a biological, pharmacological or pathophysiological basis for the observation is just an observation and needs further evaluation concerning the underlying processes.

Non-linearity of toxicokinetics is a known cause of non-monotonic-dose-response (NMDR) if, for example, the mode of action is concentration dependent (for example, two receptors with different actions and different K_Ds). Further underlying causes of an absence of a monotonic dose-response may be time-dependency with receptor down regulation, induction/inhibition of metabolizing enzymes, changing responses in the chain of events from the cellular level to the final observed effect by "adaptive" responses, or "compensatory" pathways.

Biologically-based models are built using physiological knowledge about body composition, blood flows, basic mechanisms of distribution, metabolism, excretion (chemical independent) and about the events, pathways and regulatory mechanisms at the cellular, tissue / organ and whole organism level. "Prior" information is used to build the models. Whereas biologically based kinetic models are well developed, there is a lack of detailed knowledge on important elements required for the construction of biologically based dynamic models (also known as computational systems biology pathway models). The construction of dynamics models has some overlap with the construction of models in systems biology. For the disruption of the system by external stimuli we face additional modelling problems as the dose-response can be determined by different mode of actions (MoA), composite endpoints (more than one mode of action), and (counter) - regulation at the different levels, to mention some of the factors.

When we discuss so-called low dose effect (effects which are observed at low doses, near human relevant levels, but not at higher doses) important questions must be answered. The most important is what is the relevance of low dose effects for the human population? Biological plausibility and knowledge on the likely mode of action are prerequisites for using the information in risk assessment. In order to be able to elucidate the question on the mode of action, or the underlying biological mechanism we need appropriate tools. Cellular systems, modified to address a specific problem (e.g., expressing nuclear receptors) might be helpful in this respect. However, we should be careful because *in vitro* test conditions might also be an underlying cause for non-monotonicity.

3.2.2. How to implement variability in toxicokinetics and toxicodynamics, and critical time windows of exposure / susceptibility, in dose-response modelling and hazard characterisation

Based on currently available data it is difficult to implement variability into dose-response analysis. Individual, non-aggregate animal data are therefore required, but are often not available. However, if we can assume a distribution type, Monte Carlo simulation is a means of simulating variability. Suitable kinetic data may be available, and in such cases simulations have been performed.
Critical time windows are covered by the existing testing paradigms and standard tests in animal testing. For example, the extended 1-generation study with a sufficient number of animals per group may be enriched, and it encompasses, for example, methylation to cover epigenetic mechanisms. It should, however, be mentioned that it is not straightforward and requires specialist expertise and particular caution is required when extrapolating the windows of exposure in development in animal models to windows of exposure in human development.

If a "low dose effect" is observed, the logical answer of the investigator should be that the dose range for testing should be expanded to test doses which are below the dose at which the low dose effect is observed. *In vitro* results could inform the dose selection for the *in vivo* study.

3.2.3. Effects of routes of exposure on toxicokinetic and toxicodynamic processes

The current paradigm is based on the assumption that the concentration at the site of action is the input for the toxicodynamic processes. Hence, for extrapolating from one route of exposure to another route of exposure physiologically based toxicokinetic (PBTK) modelling is instrumental, because it allows modelling of the concentration time profile for different routes of exposure. However, it is necessary to build the model on the information relevant for the chemical or chemical mixture under consideration. For example, if metabolism at the site of entry (e.g., in gut wall / gut microflora) produces the toxicant, we have to expect that different effects may occur when the exposure is by the oral as compared to the dermal or inhalation route of exposure. There is yet not much evidence on the effect of different routes of exposure for low dose effects (e.g. mycotoxins are effective at very low doses via the oral route, but mycotoxin exposure by dust inhalation may also be relevant).

3.2.4. Integration of in vitro effects to in vivo whole body response

Integration of *in vitro* effects into a model of *in vivo* body response is presently still a goal. Some examples are published where the authors extrapolated the concentrations used in *in vitro* cultures into external exposure (doses) *in vivo*. The basis for the extrapolation is to rely on measured concentrations time profiles in the *in vitro* study, so-called biokinetic studies, rather than to use nominal concentrations.

Related to toxicodynamics, low dose effects and non-monotonic responses can be observed in transcriptomics / proteomics / metabolomics. The present status of knowledge allows classification (qualitative responses) of chemicals– e.g., as genotoxic/non-genotoxic/hepatotoxic chemicals– based on pathway analysis and principal component analysis. Before we can use this information on a quantitative level we must accumulate more knowledge. For some endpoints (e.g., reproductive toxicity) we have a whole array of *in*

vitro models. It is necessary to integrate the results of the different models into a system allowing description of the dose/concentration response profile.

If we see a NMDRC in the *in vitro* study we must know the mode of action to explain the observation at the biological level to draw further conclusions. The observation of a NMDRC *in vitro* only would warrant but not necessarily trigger the conduct of *in vivo* studies.

3.2.5. Physiologically-based models in dose-response assessment

For kinetics we have PBTK models which can be used to investigate non-monotonic dose response behaviour. Several publications describe the principles to construct models and their application in dose-response modelling and in route-to-route extrapolation. One further application is the retrospective reconstruction of exposure, which is often only applicable to chemicals with long half-lives. For dynamics most of the models used are empirical models fit to the observed data. They have the limitation that, in principle, extrapolation outside the range of the observation is accompanied by high uncertainty. To be able to construct physiologically-based toxicodynamic models for dose-response analysis we need to assemble appropriate data. It is expected that support will be given by systems biology, as well as by other approaches that use real human data. Several groups make every effort to describe the physiology of organs and tissues to arrive at a mechanistically relevant description of the full body ("physiom"). It is the hope that their data will form the base for construction of physiologically-based toxicodynamic models.

Additional questions raised and discussed

Which dose-response curve should be expected with specific types of substances??Should the dose spacing be changed in these instances? Are there areas in biology where you can expect U-shaped effect curves?

These questions have been discussed by participants in the breakout group, without clear answers. It was again mentioned that feedback mechanisms and compensatory effects have to be considered. Compensatory effects can be seen in toxicodynamics. Examples are receptor down regulation and counter-regulation, which is used by the body to maintain the system in equilibrium. Some participants stated that NMDRC should not be disregarded in risk assessment, whereas others underscored the necessity to understand the mode of action before drawing conclusions for risk assessment.

3.3. Discussion Group 3 - Low dose effects: Is there sufficient evidence for non-monotonic dose-response curves?

Chair: Paul Brantom – Rapporteur: Christophe Rousselle

Recently, the evaluation of low dose effects of chemicals has been discussed (Birnbaum, 2012) as well as NMDR such as "hormesis" (inverted U-shaped or a J-shaped dose–response curve) in which opposite effects have been observed at low, compared to high, doses for the same measured parameter (Connolly and Lutz, 2004; Calabrese and Blain., 2011; Vandenberg et al., 2012).

The so-called "low dose" hypothesis dates back to the late 1990s, on the basis of studies claiming that hormonally active environmental agents can cause a variety of effects, mainly reproductive and developmental, at "low doses" (vom Saal and Sheehan, 1998). Low dose

effects have been suggested for a number of chemicals that mimic natural hormones, such as some pesticides, dioxins, polychlorinated biphenyls (PCBs), polychlorinated dibenzofurans (PCDFs) and bisphenol A (BPA).

A dose-response curve is non-monotonic when the slope of the curve changes sign somewhere within the range of doses examined. Non monotonicity is not synonymous with low dose, because there are low dose effects that follow monotonic dose-response curves.
The consequence of non-monotonic dose-responses for toxicity testing is that a safe dose determined from high doses does not guarantee safety at lower untested doses that may be closer to current human exposure.

These two theories challenge key concepts in toxicology and risk assessment, such as the existence of a "safe" threshold dose for most (non genotoxic) chemicals, and the possibility to predict the effects of a chemical at low doses from its effects at higher doses. As yet, these claims are still controversial and the biochemical mechanisms by which these effects would occur are not well understood.

There is a great interest and debate within the scientific community concerning the scientific validity of these hypotheses and how risk assessment process may include these observations.
After a short introduction on the overall process, the chair asked all participants to respond to the following questions emphasising that it was not the aim of the discussion to arrive at a consensus but to gain understanding for the range of views.

3.3.1. Defining low dose effects and non-monotonic dose-responses; what do they mean in the context of this Colloquium?

Different definitions of "low-dose effects" include effects that occur in the typical range of human exposures, or at environmentally-relevant doses or at a dose administered to an animal that produces blood concentrations of that chemical in the range of what has been measured in the general population. "Low dose "may also be considered as doses below those used in traditional toxicological studies, or at doses below the presumed NO(A)EL or BMDL expected by the traditional testing paradigm (Melnick et al., 2002; Vandenberg et al., 2012).

The definition of NMDR was less controversial among the participants: a dose-response curve is non-monotonic when the slope of the curve changes sign somewhere within the range of doses examined.

When defining effects at low dose by comparison with the NOAEL derived from standard toxicological studies it is implicit that the effects observed at these low doses are well characterised and considered as adverse. For some compounds, based on results from *in vitro* or *in silico* testing (endocrine activity from *in vitro* binding, transcriptional activation, etc.), some "low dose" effects may be expected on end-points not included in classical regulatory toxicology studies (e.g. effects mediated by hormonal disturbance) and need to be investigated in a dedicated study. In this case, the "current" NOAEL, which refers to the NOAEL derived from "high" dose standard tests, should be the upper limit for selecting dose ranges for this new study required to address low dose concerns. This kind of experiment should also consider particular windows of susceptibility (e.g. prenatal exposure).
In a regulatory context, when establishing reference values considered safe for the exposed population (Tolerable Daily Intake, Occupational Exposure Level, etc.) which are derived from a point of departure (NOAEL / LOAEL / BMDL) divided by uncertainty factors, if a

NMDR relationship cannot be excluded, based on the available data, some participants consider that this so called "safe level" should be experimentally tested to confirm its safety.

As science is moving fast, levels of low doses may also change from time to time, following the development of new tools and methods. It is then important to consider low dose effects in relation to the method for investigating them and its sensitivity.

As it was recognised that non monotonicity is not synonymous with low dose, because there are low dose effects that follow monotonic dose-response curves and *vice versa*, non monotonicity below the "current" NOAEL was not considered as a requirement for considering low dose effects.

3.3.2. Is the current scientific evidence for low-dose effects and non-monotonic dose-responses for endocrine-active chemicals convincing? (in vitro, in vivo / in mammalian species and epidemiological evidence)

First, the group noted that the issue of low-dose effects and non-monotonic dose-responses should not be considered only in the context of endocrine disruption but should include other types of effects not mediated by endocrine pathways.

It was acknowledged by the discussion group that NMDR and low dose effects have been described for certain substances and are credible. It was commented that there is good evidence from experimental data (e.g. those showing ED/repro/receptor-mediated effects) but there is still a need for more epidemiological evidence. For around half of the participants in the discussion group, the currently available evidence is rather convincing.
The participants recognised that the quality of data should be assessed for studies showing NMDR as for any other studies. The statistical evidence and mechanistic plausibility of NMDR should be assessed before concluding that it is a NMDR. Although biological plausibility is important it was noted by some participants that one cannot exclude NMDR even if we do not know at this moment the mechanism for such effects.

It was also acknowledged that there is good evidence of NMDR for some types of adverse effects, such as sex ratio, sexual behaviours, uterus weight, etc. In the context of ED, MDR as well as NMDR can be observed with the same compound, on the same target organ, depending on the mode of action. Not all endocrine disruptor effects show a NMDR.

It was also discussed if there is any evidence of NMDR outside those effects implying receptor interactions or more generally protein binding? No other examples were identified, except maybe in the radiation field. But it was also recognised by the participants that excluding a receptor-mediated or protein-binding mediated effect for a compound is not an easy task. In this context, a new acronym was proposed by some participants: Neuro-Immuno-Endocrine System which may cover this kind of NMDR.

3.3.3. If not, which data are necessary to provide conclusive scientific evidence for the occurrence of low dose effects and non-monotonic dose-response curves?

To confirm low dose effects or NMDR curves, some of the group proposed that different species of test animals should be tested, the results should be reproducible and if possible, the mode of action explained. Then human relevance of the observed effects should be considered.

The participants acknowledged that investigating these kinds of effect requires the use of best methods in terms of sensitivity. QSAR and High Throughput Screening methods can be used to test receptor interactions which may give hints for low dose effects or NMDR curves. However, even if new methods are currently available in laboratories (e.g. pharmaceutical laboratories), they need standardisation and should be used under well controlled conditions.

Some participants considered that seven doses could be the optimal number to investigate NMDR. In the context of low doses, all recognised that a study with only one dose cannot be used alone in a risk assessment process, but can be used if supported by other studies or to refine mechanisms of an effect previously demonstrated after testing other doses. However, it was discussed that animal welfare should be considered: if more dose groups are to be used, The BMD approach might be preferred and allow a reduction in the number of animals per dose group.

The group discussed what should be required for all unknown or novel compounds compared to a particular one for which we already have some concerns. The participants recognised that it is a case by case approach but it is still important to have some generic requirements for a new compound.

As regulatory toxicology is mainly driven by the OECD framework, the participants made some proposals on how to improve OECD guidelines. Participants found that OECD guidelines are not always comprehensive, being mainly end-point based and not including enough mechanistic explanation. The participants recommended that new end-points not investigated in the current OECD guidelines but identified in academic studies should be considered for future inclusion. Some participants were also in favour of a new OECD guideline for an experimental study with an *in utero* exposure and follow up of the F1 generation for most of the life-span. This kind of protocol may be used to investigate for example carcinogenic effects occurring later after an exposure during the developmental period. Requirement for toxicokinetics data could also be a further way to help bridge between the observed effects to internal doses, which could then be compared to exposure levels and human bio-monitoring data.

Concerning non GLP / OECD studies, the group considered that criteria should be established to assess this kind of study in a regulatory context.

Some participants proposed to test compounds in real conditions (e.g. contaminants in food or ingredients in formulations) but the group recognised that this would be very difficult, and from past experience unlikely to yield useful data.

3.3.4. Are the current testing paradigms adequate to detect "low dose effects"? If not, how experimental design could be improved to address properly low-dose effects and non-monotonic dose-responses?

This question was rephrased by the group to focus on the issue "How could the new information based on low dose effects and NMDR be used in risk assessment?"
The participants acknowledged that in a practical way, it would be almost impossible to experimentally test a large number of doses for a new chemical covering all exposures from the very lowest level to those occurring in the occupational environment. The Benchmark Dose approach which is more robust when increasing the number of doses even with less

animals per dose could in this context be an interesting alternative approach to the classical NOAEL / LOAEL derivation, although application of this method to NMDR is untried.

The participants considered that it is important to have a step by step approach to identify compounds for which insufficient information is available. *In vitro* or *in silico* tests could be performed as a first step to screen compounds for which additional *in vivo* data are required. All suspected endocrine disrupters or endocrine active chemicals should be screened at least *in vitro*. Even "old" chemicals, already used for many years, should be assessed in respect of low dose effects and NMDR curves. But the question is then, who will be responsible for/ take responsibility for carrying out these new assessments?

The participants also recognised that chemicals for which there is exposure of susceptible individuals, deserve particular attention and should be tested in an appropriate way.

The participants discussed the need for experts and risk assessors to have access to the raw data when using studies published in the scientific literature. It is indeed often not sufficient to rely on the compiled data reported in the manuscript and individual results are useful to evaluate for example the statistical plausibility of a NMDR. Badly reported studies are often rejected in a risk assessment process even if the raw data could have been useful. The participants recognised that it is often difficult to get these data by just asking the authors for them. To avoid this loss of knowledge, it could be required by the editor of the scientific journals, before publishing the study, to make the raw data available, for example on a dedicated website as supplementary information. This could be included in the reviewing process. One participant mentioned the ARRIVE guideline[3] (Animal Research: Reporting *In vivo* Experiments, 2010). This guideline provides a checklist for those preparing or reviewing a manuscript intended for publication. It could be a good place to make the recommendations to make raw data available.

3.3.5. How to model non-monotonic dose-response in the context of a quantitative risk assessment?

If for a compound, NMDR is observed and considered reliable, then the question is: "How can the risk be assessed?". Is it still possible to derive a NOAEL / LOAEL or should other approaches be considered? If so, which one(s)?

Participants were aware that modelling curves depend first on consideration of what is an adverse effect. In the context of endocrine disruption, the endogenous levels of hormones should be taken into account as well as the natural components in food with similar effects.

Some participants recommend modelling the first stage of a NMDR curve or the portion representative of the exposure level for the targeted population. For a U-shaped curve, participants acknowledged that an infinite effect at the very low dose is not plausible and that effects at low dose should be compared to the control group or to a range of normal values in an unexposed population, to decide where to put the "low" NOAEL.

Some participants proposed a non threshold approach to characterise the risk of a compound showing a NMDR. An extra risk could then been calculated for a unit of dose exposure. However some participants warned about the regulatory consequences, in the light of the preceding discussion that an infinite effect is not plausible.

[3] Available from http://www.nc3rs.org.uk/downloaddoc.asp?id=1206&page=1357&skin=0

Except for these few proposals, the group recognised that this issue is still under debate and, at the present time, no agreed methodology has been developed to deal with NMDR curves for assessing the risk. It will be a challenge for risk assessors for the coming months and probably years.

3.4. Discussion Group 4 - Impact for risk assessment

Chair: Anthony Hardy – Rapporteur: Fernando Aguilar

Impact for risk assessment of the low dose / non-monotonic response was intensively discussed among the participants. For the needs of the discussion participants agreed to assume the validity of the low dose / NMDRC hypothesis and that an effect which is observed at low dose ranges is adverse. The group pointed out, however, that the outcomes from ongoing debates in the other discussion groups in this colloquium would have an influence on its observations (e.g. what is an adverse effect at low dose range?).

3.4.1. Assuming a general acceptance of the scientific validity of the low dose / non-monotonic dose-response curve hypothesis, does this dictate a need for new risk assessment approaches?

Overall, participants considered that the existing risk assessment paradigm is applicable to assess risk that could be associated with low dose / non-monotonic responses. The group considered that the identification of hazard for substances showing non-monotonic response could be approached by using a "classical" read-across → *in vitro* testing → *in vivo* testing. However, some adjustments would be needed to take into account particularities of the low dose / non-monotonic responses. For example, in some non-monotonic responses, the lack of data between the lower dose that shows an effect and the dose not showing an effect would not allow identification of a no-observed effect level (NOEL) or a no-observed adverse effect level (NOAEL). The group observed that in non-monotonic responses, different mechanisms might be at work, further complicating the possibility to identify relevant NOELs or NOAELs. This would be particularly important for substances with few toxicity data which, for example, might undermine the ability to fine-tune the assessment approach during hazard characterisation. However, the group recognised that in the case of non-monotonic responses in substances with much more data available, the need for fine-tuning could also be hampered by the fact that existing data has been gathered mainly at high dose ranges.

The group considered that in terms of animal welfare and testing a systematic exploration of low dose-responses would not be possible. However, it was observed that data available on regulated chemicals could be used to reduce the need for *in vivo* testing, from example from the TOX 21 project[4] which is screening thousands of chemicals to predict their potential toxicity, including at low dose ranges. The increased analysis of such data might be useful to identify particular chemicals that need more precise investigation at low dose ranges and the group considered that some tools to do this type of analysis are already available. However, it was also noted that information on effects of mixtures of substances, with similar or dissimilar modes of action, is lacking and that additional tools might be needed to address this situation including: a) tools to take into account kinetic and dynamic parameters to better identify internal doses, b) adapted QSAR tools to better integrate data from existing data, such as TOX 21, c) validated follow-up quantitative approaches to take into account toxicokinetic

[4] Available from http://epa.gov/ncct/Tox21/

studies *in vivo* for example, and d) convergent approaches to identify what is an adverse effect when dealing with low-dose-responses.

3.4.2. Are different approaches already in use in risk assessment appropriate to deal with low-dose effects and non-monotonic dose-response curves (e.g. is there any need for additional uncertainty factors, does the Margin of Exposure approach covers these responses, can the TCC concept be applied to these responses), if not which data gaps would need to be filled to achieve a full risk assessment of this type of compound?

The group discussed the possibility of applying specific risk assessment approaches to assess low-dose / non-monotonic responses such as introducing additional uncertainty factors (UF), applying Margin of Exposure (MoE) or the Threshold of Toxicological Concern (TTC). Introducing additional UF was not considered feasible when dealing with low-dose / non-monotonic responses since default UF (for example 10 x 10) are derived from studies conducted at high dose ranges and therefore extrapolation of default factors to low dose ranges would not be appropriate. Concerning the TTC the group noted that the EFSA opinion on this matter states that if there are data showing that a substance has endocrine activity, but the human relevance is unclear, then these data should be taken into consideration, case-by-case, in deciding whether or not to apply the TTC approach. If there are data showing that a substance has endocrine-mediated adverse effects, then, as would be the case for adverse data on any other endpoint, the risk assessment should be based on the data, rather than the TTC approach. Unfortunately, there was insufficient time to discuss the applicability of MoE to low-dose / non-monotonic responses. Overall, the group considered that the current paradigm and UFs applied routinely appear sufficient to assess risks for the general population. The group recognised nonetheless that it is necessary to consider further the need for revised strategies to assess data that point sufficiently strongly to the existence of a non-monotonic response.

Concerning the exposure assessment of substances showing non-monotonic responses, the group considered that existing exposure assessments, that are crude but over-protective, already may cover exposure to these substances. One particular active discussion point was the fact that a tiered exposure approach, which defines the quantity of toxicity data needed to do a risk assessment of a substance, does not take into account low-dose-responses. Food contact materials were cited as an example for which according to existing EFSA guidelines, the amount of toxicity data needed is linked to migration rates and consequently to exposure estimations from packaged food. In the case of a substance associated with a low migration rate, toxicity data would be minimal, and particular studies would not be required given the anticipated low exposures. This latter point leads to the observation that available information on non-monotonic curves could be used to better define the range of doses to be screened for potential low-dose-responses.

Therefore, it was stressed that more knowledge on low-dose-responses is needed to take into account several factors such as the agonistic and antagonistic effects of mixtures, which may compensate each other, as well as kinetic data to better understand and estimate the internal doses of substances showing non-monotonic responses. The group suggested that all available information should be integrated to apply risk assessment to identify other potential end-points that could be taken into account within the existing risk assessment procedures. It was pointed out that risk characterisation of substances showing non-monotonic responses such as endocrine active substances needs revision.

3.4.3. Assuming a general acceptance of the scientific validity of the low dose / non-monotonic dose-response curve hypothesis, how to take critical windows of susceptibility into account in the risk assessment process of these compounds?

The group discussed how to take into account windows of susceptibility in the risk assessment process of substances showing low-dose / non-monotonic responses and considered that existing toxicity methods have been improved to take into account potential windows of susceptibility for low dose-responses. Whilst acknowledging that at present particular windows of susceptibility are not routinely included in testing protocols, the group noted that a requirement for such testing would imply that many chemicals could have such effects, which may not be the case in reality. Furthermore, in the case of substances with limited toxicity data, this type of effect may not be identified. The group considered it more feasible to identify "signals" suggesting windows of susceptibility at low doses. Gathering data on such signals might allow making more informed decisions. As more science on these responses becomes available, knowledge would increase and the risk assessment approach would evolve too.

3.4.4. Can traditional "gold-standard" toxicology studies be coupled to targeted endpoint research studies to derive health-based guidance values for this type of compound?

Finally, the group discussed "gold standard" studies and the possibility that they can be coupled to targeted endpoint research studies to derive health-based guidance values for low-dose / non-monotonic responses. The group considered that the first parameter to be taken into account is how the quality of research studies has been assessed (e.g. quality assured, guideline compliant studies). If quality is assured, then such research studies, if they integrate other endpoints, could be used to derive health-based values. However, the group also discussed the "validity" of the endpoints chosen since it was noted that test methods are increasingly sensitive thus increasing the capacity to identify hazard. The important question then is does the newly identify hazard matter for risk assessment?

4. FINAL PLENARY DISCUSSION AND CONCLUSIONS

The final session was dedicated to reports back from each of the four discussion groups, in the form of a presentation by the rapporteur followed by debate on the outcomes of each discussion group. This was followed by a general discussion where conclusions and recommendations of the colloquium were discussed.

The low-dose effect and non-monotonicity hypotheses challenge key concepts in toxicology and risk assessment, and also the possibility to predict the effects of a chemical at low levels of exposure from its effects at higher levels of exposure. This colloquium aimed to exchange views on the topics of low-dose effects and NMDRC and how these phenomena should impact the current toxicological risk assessment paradigm. It was not the intention to reach consensus on the scientific acceptability or credibility of these concepts. For the purpose of constructive discussion, participants in certain discussion groups were asked to assume the validity of the low-dose/NMDRC hypotheses as a starting point.

It was acknowledged during the colloquium that NMDRC and low dose effects have been described for certain substances and are considered credible by part of the participants in the Colloquium. It was the view of these participants that there is evidence from experimental data for such effects. The epidemiological evidence is however very limited and more work is needed.

It was stated that the quality of data for studies showing NMDRC should be assessed as for any other studies. The statistical evidence and mechanistic plausibility of NMDRC should be analysed before concluding that a dose-response is non-monotonic. Although biological plausibility is important it was noted by some participants that one cannot exclude NMDRC even if the mechanism for such effects is not known at this moment. However, other participants considered that mechanistic understanding is a prerequisite in order to render the results relevant for risk assessment

It was also discussed if there is any evidence of NMDR outside those effects involving receptor interactions or more generally protein binding. No other examples were identified, except possibly in the radiation field. But it was also recognised by the participants that excluding a receptor-mediated or protein-binding mediated effect for a compound is not an easy task.

The **adversity** of "low-dose effects" or effects for which NMDRCs are reported, was an important topic of discussion. A definition of adversity was presented stating: "adverse effects are biochemical, morphological or physiological changes (in response to a stimulus, in this case the chemical substance) that either singly or in combination adversely affect the performance of the whole organism or reduce the organism's ability to respond to an additional environmental challenge". Participants considered that this definition does not need to be different for low dose effects or NMDRCs. In identifying adversity it is however a scientific challenge to determine what level of change in a biological parameter is of toxicological relevance, in particular if the measurement outcome has a high variability. In addition, it has to be agreed what the sensitivity of a study should be, or in other words what statistical power is appropriate given the size of the toxicologically relevant effect. When planning a study the magnitude of the effect under consideration and the statistical power are the determinants for the number of animals.

It was considered by the meeting that the available toxicity testing protocols in animals for identifying the hazards of chemicals can be used to study the effects of chemicals at low doses or to describe NMDRCs. However, the number of dose groups has to be increased especially in the low dose area taking into account human exposure levels. When an increased numbers of doses are tested, the number of animals used will increase accordingly.

In vitro studies can be used for priority setting and for identification of modes of action, but are normally not useful to define adversity of effects. The same holds true for changes in genomic and proteomic responses. In contrast, results from epidemiological studies may be useful to identify adverse effects at low doses, but the causality is difficult to prove, mainly because of the retrospective nature of exposure assessment. Participants agreed that results of studies demonstrating low-dose effects and/or NMDRCs need to be reproducible, whether these are *in vitro, in vivo* studies in animals or human studies.

The meeting considered that in order to **establish a PoD** from NMDRCs well-described dose-response curves are needed, with more doses tested in the low dose range than what is currently common practice. At present, since the whole dose-response curve very often is not known, there will be uncertainties in defining a NOAEL for a substance with a suspected NMDRC. The lowest dose tested should be based on estimated human exposure.

Where low-dose effects or NMDRCs have been reported, there is no standardisation of the toxicological end-points investigated, and these are determined case by case dependent on the suspected mode of action of the chemical being tested . However, how the term "mode of action" should be interpreted was not a topic in this colloquium. In many areas, validated studies are not available and the assessment may need to rely on non-validated studies based on appropriate experimental design and reporting. Experiments from new non-validated models can be used when they are properly described, variability is understood and controlled, and when reproducibility and reliability is established.

The profile of a dose-response curve is determined by both the **toxicokinetics and the toxicodynamics** of the substances. Hence, non-linearities in the **kinetics**, which might be dose dependent can influence the shape of the dose-response relationship. For toxicokinetics, physiologically based models are available which can be used to model the fate of a substance. Several publications describe the principles to construct models and their application in dose-response modelling and in route-to-route extrapolation. Such models may be used to gain a better understanding and estimation of internal doses of substances showing non-monotonic responses.

In contrast, data required for the construction of **biological models and the related dynamic processes** are generally not available. Current modelling of dose-response curves (**dynamics**) is commonly done by fitting a curve to the observed toxicity response data. Such models have the limitation that extrapolation outside the range of the observation is accompanied by uncertainty and in addition, the fitted function is a mathematical expression rather than a description of the underlying biological mechanisms.

The actual dose-response can be influenced by different MoAs along the dose-range, composite endpoints (more than one mode of action), different feedback loops at the different levels, to mention some of the factors. The biological plausibility of the effect must be considered and some knowledge of the mode of action is needed when using the information on "low-dose effects" or NMDRCs in risk assessment.

Integration of *in vitro* effects into a model for description of *in vivo* responses is presently not feasible, but some authors have presented examples in which the concentrations used in *in vitro* cultures are extrapolated to external exposure (doses) *in vivo*. If in an *in vitro* study an NMDRC is observed, the MoA has to be known in order to explain the observation at the biological level and to draw further conclusions. Low dose effects and non-monotonic responses can be observed in "-omics" studies which could be helpful to clarify MoAs. However, there is no possibility, yet, to include this information into risk assessment strategies in a quantitative way. In absence of information of the MoA, an observation of an NMDRC *in vitro* would not necessarily trigger *in vivo* studies.

With respect to the **impact on risk assessment** of low dose effects it was considered that the existing paradigm is applicable to assess risk associated with NMDRCs. Identification of hazards for substances showing NMDRC could be approached by using a "classical" read-across → *in vitro* testing → *in vivo* testing scheme. However, some adjustments would be needed to take into account particularities of the low dose/non-monotonic responses. For derivation of PoDs for risk assessment, some adjustments may be needed. Lack of data points between the low dose that shows an effect and the putative dose without an effect (i.e. even further down the dose range) would hamper identification of a NOEL or a NOAEL. Also for data-rich substances for which an NMDRC is anticipated, fine-tuning could be hampered by the fact that existing data has been gathered mainly at the high end of the dose range. In non-monotonic responses, different mechanisms might be at work for different parts of the dose-response curve, complicating the possibility to identify relevant NOELs or NOAELs.

Extensive studies of low dose-responses would be demanding in terms of numbers of animals. Therefore it is a goal to identify less cost-intensive in vitro assays. The remark was made that data available on regulated chemicals, for example from high throughput screening assays, might be used to predict potential toxicity, including toxicity at low dose ranges. Analysis of such data might be useful to identify chemicals that need more precise analysis at low dose ranges.

Information on **effects of mixtures** of substances, with similar or diverging modes of action is lacking. It was stressed that additional tools might be needed to address low-dose-responses of mixtures, to take into account several factors such as the agonist and antagonist effects of mixture components, which may compensate each other.

Introducing **additional UF** was not considered feasible when dealing with NMDRC since UF are developed for conventional toxicity studies, which use relatively high levels of exposure and which are analysed assuming a monotonic dose-response relationship. Therefore, extrapolation with linear default factors would not be appropriate. Concerning the **TTC concept**, the group noted that the EFSA opinion on this matter states that if there are data showing that a substance has endocrine activity, but the human relevance is unclear, then these data should be taken into consideration, case-by-case, in deciding whether or not to apply the TTC approach. If there are data showing that a substance has endocrine-mediated adverse effects, then, as would be the case for adverse data on any other endpoint, the risk assessment should be based on the data, rather than the TTC approach. It was recognised that it is necessary to discuss on the need for revised strategies to assess data that point sufficiently strongly to the existence of NMDRCs.

It was recognised that tiered testing approaches, which define the quantity of toxicity data needed to do a risk assessment of a substance, do not take into account low-dose or non-

monotonic responses. Available information on non-monotonic curves could be used to better define the range of doses to be screened for potential low-dose-responses.

Another part of the discussion dealt with how to take into account **windows of susceptibility**. Existing toxicity methods have been improved to take into account potential windows of susceptibility for low dose-responses. It was mentioned that there is a challenge in extrapolating windows of exposure in developmental studies in animals to windows of exposure in human development. Routine testing for effects linked to windows of susceptibility implies that many chemicals could have such effects, which may not be the case in reality. Furthermore, in the case of substances with limited toxicity data, this type of effect might not be identified. It would be more feasible to identify "signals" suggesting effects with windows of susceptibility at low doses. Gathering data on such signals might allow making decisions that are more informed and as more science on these responses becomes available, knowledge would evolve and risk assessment approaches will have to evolve too.

It was considered that "gold standard" studies can be coupled to **targeted endpoint research studies** to derive health-based guidance values for low-dose / non-monotonic responses. However, the quality of such targeted endpoint research studies must be addressed. If quality is assured then such studies could be used to derive health-based values for substances displaying NMRDCs, if they integrate other endpoints. However, the "validity" of the endpoints chosen should also be addressed.

Conclusions

It should be noted that no extensive discussion was conducted on the question of whether there was sufficient scientific evidence for the existence of "low-dose effects" and / or NMDRCs. However, as indicated previously, for the purpose of constructive discussion participants in the breakout groups were asked to assume the validity of the low-dose/NMDRC hypotheses as a starting point. The following were the main conclusions of the meeting:

- An adequate and generally accepted definition of "low-dose effects" and of NMDRC is needed in order to facilitate discussions.

- The amount of evidence needed to decide if in a particular case a "low-dose effect" or an NMDRC has to be taken into account should be defined.

- Information may be obtained from *in vitro* and *in vivo* studies to determine biological plausibility.

- Data on toxicokinetics, MoA and toxicodynamics will be helpful to understand the nature of the observations and to link internal dose estimates to occurrence of adverse effects.

- The criteria for adversity should be the same for all types of effects.

- It should be possible to derive Points of Departure (PoDs, NOAEL / BMDL) for risk assessment in studies with an adequate (extended range) number of dose levels, in particular in the lower dose range and even if there is a NMDR.

- Information should be obtained from well-designed studies covering wide dose ranges with more than usual dose groups and sufficient animals per group.

- Dose selection may be based on observations in epidemiological studies or on estimates of human exposure to cover the low exposure ranges more adequately.

- It was noted that although the established principles of toxicological risk assessment would still be applicable, adaptation of these techniques might be needed.

- It was generally considered that tiered approaches for hazard assessment guided by exposure estimates might not be adequate for substances for which an NMDRC is suspected.

Overall, participants considered that the existing risk assessment paradigm is applicable to assess risks that could be associated with low dose / non-monotonic responses. Some participants stated that NMDRC should not be disregarded in risk assessment, whereas others underscored the necessity to understand the mode of action before drawing conclusions for risk assessment. Thus, implementation of "low-dose effects" and NMDRCs in risk assessment strategies presents a scientific challenge and development of intelligent testing strategies to deal with these phenomena is necessary.

It was clear that different views on the significance of "low-dose effects" and NMDRCs might circulate in different scientific disciplines. Assuming that low-dose effects and NMDRCs are to be accepted as a "fact-of-life", it should be decided whether these are applicable for specific MoA, or whether they are universal principles applicable to any MoA.

From the discussions, it became clear that there is a need for an in-depth analysis of available studies in which these phenomena have been reported. It was recommended that as a follow-up, EFSA should consider to set up an *ad hoc* multidisciplinary working group to examine the scientific evidence for 'low-dose effects" and NMDRCs, and for which MoAs they are applicable.

5. ABBREVIATIONS

ADI	Acceptable Daily Intake
ARRIVE	Animal Research: Reporting In Vivo Experiments
BMD	Benchmark Dose
BMDL	Benchmark Dose Lower bound
BPA	Bisphenol A
EC	European Commission
ECHA	European Chemical Agency
ED	Endocrine Disruptors
EDC	Endocrine Disrupting Chemicals
EFSA	European Food Safety Authority
EPA	U.S. Environmental Protection Agency
EU	European Union
FDA	U.S. Food and Drug Administration
IFT	U.S. Institute of Food Technologists
IPCS	International Programme on Chemical Safety
JRC	EC Joint Research Center
KDs	Dissociation Constants
LOAEL	Lowest Observed Adverse Effect Level
MoA	Mode of Action
MoE	Margin of Exposure
NMDR	Non-Monotonic Dose-response
NMDRC	Non-Monotonic Dose-response Curves
NOEL	No-Observed Effect Level
NOAEL	No-Observed-Adverse-Effect-Level
OECD	Organisation for Economic Co-operation and Development
PBPK	Physiologically based pharmacokinetic
PCBs	Polychlorinated biphenyls
PCDFs	Polychlorinated dibenzofurans
PoD	Point Of Departure
QSAR	Quantitative Structure Activity Relationship
REACH	EC Registration, Evaluation, Authorisation and Restriction of Chemical substances
SCENIHR	EC Scientific Committee on Emerging and Newly Identified Health Risks
TDI	Tolerable Daily Intake
TTC	Threshold of Toxicological Concern
UF	Uncertainty Factors
WHO	World Health Organization