SCIENTIFIC COMMITTEE AND EMERGING RISKS UNIT

EFSA Scientific Committee Opinion on biological plausibility of non-monotonic dose responses and their impact on the risk assessment

Table of Contents

1. Introduction

Even though there is no full agreement on a definition of non-monotonicity, it is generally accepted that in a non-monotonic dose-response curve, the slope changes sign (change of direction) at least once (See figure 1). This could be determined by non-linearity in toxicokinetics, toxicodynamics, or both. In principle, non-monotonicity may occur in different regions of the dose-response curve (e.g. Hill et al. 2018). Non-monotonicity occurring at the lower end of the dose-response has been often referred to as low dose effects[1]. Several studies have reported non-monotonic dose-response curves for a number of chemicals, including pesticides, polychlorinated biphenyl (PCBs), dioxins and food contact materials such as bisphenol A (BPA) and phthalates, mainly regarding their endocrine activity (EFSA 2012).

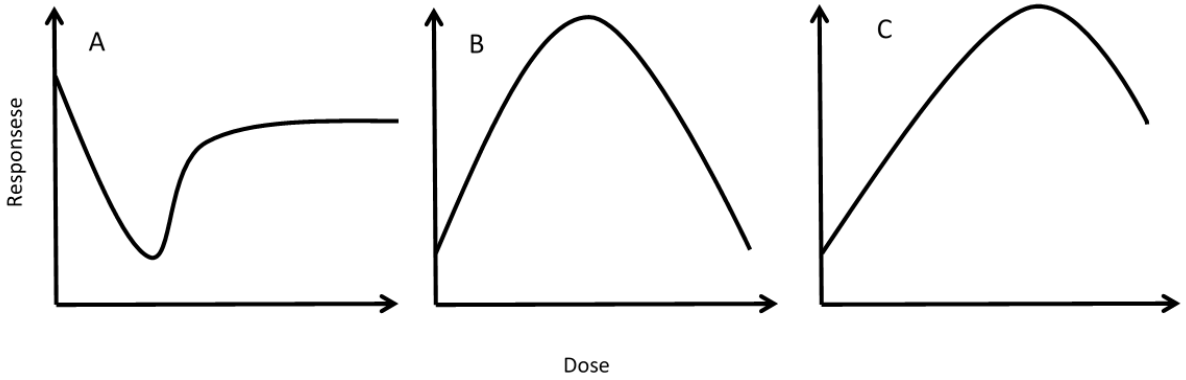


Figure 1: Examples of non-monotonic dose-response. The left figure (A) is one example of non-monotonicity occurring at the lower end of the dose-response. The middle figure (B) is an example of inverted U-shape dose-response, while the figure to the right (C) gives an example of non-monotonicity occurring at the higher end of the dose-response

Concepts describing NMDR have been described and disputed in the literature over several years. These include the concept of "hormesis" (Calabrese and Mattson, 2017), in which opposite effects have been observed at low versus high doses. These were also described for physiological reactions, with stimulatory effects being observed at low doses, followed by inhibitory effects on the same physiological parameter at high doses (Calabrese and Mattson, 2017). Connolly and Lutz (2004) described examples of non-monotonic dose-response relations that they considered as superimposition of monotonic dose-responses of components of the respective biological system.

To discuss issues around low-dose effects and non-monotonic dose-response and their potential impact on risk assessment, EFSA organized a scientific colloquium in 2012 (EFSA, 2012). The colloquium report concluded that "*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[2] should not be disregarded in risk assessment, whereas others underscored the necessity to understand the mode of*

[1] "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. EFSA, 2012 S Colloquium

[2] NMDRC: non-monotonic dose-response curve(s)

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".

In order to address these challenges, the Colloquium participants identified the need for an in-depth analysis of available studies, looking at the strength of the evidence, and for which modes of actions of these phenomena have been reported (EFSA, 2012).

Systematic review of non-monotonic dose-responses of substances for human risk assessment

To follow up on the recommendation of the Scientific Colloquium regarding the need for an in-depth assessment of current literature, EFSA contracted out a systematic review of the existing literature where signs of non-monotonic responses had been reported. The results were published as an EFSA external report (Beausoleil et al., 2016); hereafter referred to as "the Report". In that Report the scientific evidence for such NMDRs was also assessed. The systematic review, with two experts reviewing each dataset, was performed in line with the EFSA guidance (EFSA, 2010).

The Report extracted dose-response datasets from studies having at least 5 dose groups, which were then analysed by PROAST software package. The strength of the evidence was characterised using visual/statistics-based checkpoints. For this purpose, the Report proposed to use a set of six checkpoints as a tool for evaluating the evidence of NMDR in a single dataset. These checkpoints were designed to take into account that data always contain both random and non-random sampling errors. The six "checkpoints", briefly, focus on the following questions:

1. Can the apparent NMDR be explained by random fluctuations around a horizontal dose-response (= no effect at all)?
2. Can the apparent NMDR be explained by random fluctuations around a monotone dose-response (MDR)?
3. Can the apparent NMDR be explained by one single potential outlying dose group?
4. Is the effect size in one of the directions of the NMDR smaller than 5 %?
5. Is the steepness of the dose-response curve outside the range of biologically plausible/realistic dose-response shapes?
6. Does the apparent NMDR consist of more (or less) than two directions?

When the answer to the indicated question was "no", the associated checkpoint was considered "fulfilled". The first two checkpoints were based on a statistical significance test in a dose-response analysis addressing random errors in the dataset. The other four checkpoints were evaluated based on visual inspection of the dose-response plots using the confidence intervals of each response. Evaluation of the selected datasets indicated that 6 % of the *in vivo* datasets fulfilled all six checkpoints and 20 % fulfilled five checkpoints.

In total, 202 *in vivo* datasets (from 49 studies), 311 *in vitro* datasets (from 91 studies) and 9 epidemiological / human datasets (from 2 studies) were identified. 179 *in vivo* and 13 *in vitro*

dose-response datasets were analysed[3]. For 23 *in vivo* datasets there were data limitations and these could, therefore, not be analysed. None of the datasets from epidemiological/human studies were analysed. In most of the *in vivo* datasets, it was concluded that the apparent NMDR was likely caused by a single outlying dose group. That is, in total only 10 out of the 179 *in vivo* datasets fulfilled all visual/statistics-based checkpoints, while five checkpoints were fulfilled by 36 *in vivo* datasets (corresponding to 20 %). The authors concluded that Criteria for evidence of NMDR, evaluation of data and importance for risk assessment had to be further evaluated.

Probabilistic assessment

Chevillotte et al. (2017a) re-analysed the same data and developed a probabilistic assessment method to characterize NMDR curves from experimental studies. This approach involved large scale sampling to obtain 10,000 dose-response curves equivalent to the experimental curve, and a characterization procedure based on inter-dose statistical comparisons. The study focused on demonstrating the general applicability of applying probabilistic methods to evaluate the presence of NMDR. Based on resampling, the methodology was used to generate a set of values considered, theoretically, equivalent to the original data, by different permutations the probability of NMDR was assessed. Curves were characterized as non-monotonic based on the definition that it is a "change of sign in slope somewhere in the dose range tested". Such changes of sign were characterized by the presence or absence of statistically significant differences between doses. The authors examined 122 dose-response curves with different shapes from 28 publications based on their methodology.

In a follow-up study, Chevillotte et al. (2017b) added four statistical criteria to assess the robustness of the assumption of non-monotonicity and characterize the types of curves obtained. These addressed aspects of distribution and intensity, as well as minimum and maximum confirmation. The authors considered that their approach strengthens the evidence of non-monotonicity in a statistical manner, but they stressed that the statistical plausibility assessment tool should only be applied after a biological/toxicological plausibility assessment. They also stressed that the interpretation of the probabilistic results remain a prerogative of the assessor, and that there is no predefined interpretation of such probabilistic results. The authors developed a software that is available from the authors (Chevillotte et al., 2017b)). They conclude that their method provides a probabilistic and objective characterization of the type of dose-response curve, relevant for the assessment of the likelihood of non-monotonic responses.

1.1 Background and Terms of Reference as provided by EFSA

In 2012 EFSA organised a Scientific Colloquium to debate the current state-of-the-art of low-dose effects and non-monotonic dose-responses in food and feed risk assessment. The participants identified the need for an in-depth analysis of available studies, looking at the strength of the evidence, and for which modes of actions of these phenomena have been reported. This recommendation was followed up in 2014 by EFSA who contracted out a systematic review of the literature claiming non-monotonic responses and a review of the

[3] According to the ToRs, this Scientific Opinion focuses on *in vivo* studies.

scientific evidence for such NMDRs; the strength of the evidence was characterised using visual/statistics-based checkpoints (Beausoleil et al., 2016). In this review, in total, 202 *in vivo* datasets (from 49 studies), 311 in vitro datasets (from 91 studies) and 9 epidemiological / human datasets (from 2 studies) were identified. 179 *in vivo* and 13 in vitro dose-response datasets were analysed. For 23 *in vivo* datasets there were data limitations and could, therefore, not be analysed. None of the datasets from epidemiological/human studies could be analysed. In most of the *in vivo* datasets, the apparent NMDR is likely caused by a single outlying dose group. In the end, only 10 out of the 179 *in vivo* datasets fulfilled all visual/statistics-based checkpoints (6%). Anses reviewed the same data using a different but complementary probabilistic approach (Chevillotte et al. 2017a,b). Whereas a small percentage of the eligible *in vivo* dataset suggests the statistical possibility of a NMDR, the biological relevance of the statistical findings as well as the possible impact on EFSA risk assessments was, however, not assessed.

As mentioned above, the evidence for NMDR was looked at only from a visual/statistics/probabilistic point of view. In order to complete this work, there is a need to review the biological plausibility of the identified NMDRs, especially for the *in vivo* datasets. If the NMDRs should be found biologically plausible, the impact of these endpoints showing a NMDR on EFSA risk assessments should be assessed.

A statistical deviation is not necessarily the signal of a biologically relevant response; consequently, it is important to assess if the possible statistically based NMDRs identified in the report are biologically relevant. In addition, the risk assessment process aggregates several sources and lines of evidence; an effect not detected in a particular study may be covered by other studies or assessments; if this is the case, the NMDR even if biologically relevant would not impact the risk assessment outcome. Therefore, in case a biologically plausible NMDR could be identified, EFSA should address if those effects are expected to be captured through the weight of evidence process of the current risk assessment practices.

The discussion on NMDR is mostly, albeit not exclusively, driven by the assessment of endocrine active substances. Thus, there is a connection with the ECHA/EFSA guidance for the identification of endocrine disruptors in the context of biocidal and plant protection products[4] which covers exclusively the hazard identification and, in the regulatory context, is specifically applicable to pesticides and biocides. At the international level, there are several activities ongoing but there are no internationally agreed conclusions available regarding the impact on the risk assessment process of the potential existence of NMDRs. This offers EFSA the opportunity for leading the process at EU level, keeping informed JRC, ECHA and EMA. There is also opportunity for international cooperation, in particular with OECD and FAO/WHO, national agencies such as FDA and USEPA, and academic institutions such as IUTOX, EUROTOX, the International Dose-Response Society and the Endocrine Society.

Terms of Reference

The Scientific Committee was requested to prepare a scientific opinion on the biological relevance, if any, of the apparent non-monotonic dose responses identified in the external

[4] https://efsa.onlinelibrary.wiley.com/doi/epdf/10.2903/j.efsa.2018.5311

report produced under GP/EFSA/SCER/2014/01, focussing on the *in vivo* datasets fulfilling all checkpoints of the visual/statistics-based analysis. In addition, in case biological relevant non-monotonic dose responses are identified, the SC is requested to address the possible consequences for the human health risk assessments conducted by EFSA. Specifically, the SC is requested:

1. To assess the biological relevance of the non-monotonic dose responses identified *in vivo* in the EFSA external Report (Beausoleil et al., 2016.) and the follow up probabilistic assessment (Chevillotte et al. 2017a,b), based on visual/statistics/probabilistic considerations.
2. To further analyse the non-monotonic dose-responses assessed as biologically plausible, grouping them if appropriate, and evaluate their potential link with adverse effects, considering if the response induction/increase and response inhibition/decrease should be associated to the same or to different adverse outcomes.
3. To assess the biological plausibility for opposite responses at different dose levels for the adverse effects that are pivotal for EFSA assessments and usually lead the health risk assessment outcome. This should inform the assessment of the impact of any biologically relevant endpoint showing a non-monotonic dose response *in vivo*, on EFSA risk assessment outcomes.
4. To recommend the follow up actions in case biologically relevant non-monotonic dose responses impacting the risk assessment outcomes are identified. These recommendations should propose within EFSA priorities as well as priorities for international cooperation to improve future risk assessments.

Considering the time and resource limitations, the SC is suggested to use information from the OpenFoodTox database, other EFSA assessments, and the expertise available at the SC and EFSA Panels and Units.

1.2 Interpretation of the Terms of Reference

The ToRs specify that the current Opinion should focus on the NMDR data identified in the Report (Beausoleil et al., 2016.) and the follow up probabilistic assessment (Chevillotte et al. 2017a,b). In view of the length of time since these activities were completed, a search for recent scientific literature on the topic was conducted. It should be noted that it was not possible to perform a comprehensive literature search for NMDRs, as the terms monotonic and non-monotonic are not necessarily used in describing dose-response curves. The SC is aware that there are other approaches to identify NMDR (e.g. Moser et al., 2016; ECHA/EFSA, 2018), these are not the focus of the current opinion.

Both the EFSA contracted systematic review (Beausoleil et al., 2016) and the probabilistic assessment of Chevillotte et al. (2017a) were primarily focused on statistical considerations for identifying non-monotonicity. Most toxicological studies use few dose groups, which makes statistical evaluation of non-monotonicity difficult and vulnerable to elements of chance (random fluctuation). This is not an issue in other areas of biomedical science where a sufficient number of individual observations from a near continuous exposure matrix and non-monotonicity can be evaluated with less dependency of individual observations or dose groups. Needless to say, for a single study the use of statistical considerations for determining

non-monotonicity has its limitations. Firstly, such an approach does not take into consideration the possible existence of similar findings in another independent study that would argue against a chance finding. Secondly, statistical considerations cannot address biological plausibility.

In considering biological plausibility of NMDRs, the Working Group noted that nutrients, particularly vitamins, minerals and trace elements, represent a specific case, in which an overall U-shaped curve is expected. At the lower end of the dose-response relationship, deficiency of the nutrient leads to adverse effects, whereas toxicity may occur at higher doses (IPCS, 2002; EFSA-SC draft on HBGV, 2020[5]). In such cases the NMDR is explained by two distinct but overlapping biological process, which existing risk assessment paradigms can easily address. IPCS (2002) and EFSA (2020) refer to an Acceptable Range of Oral Intake (AROI) for essential nutrients, bounded by rising risks of either deficiency, as intake declines, or toxicity as intake increases. As this is a well-known situation fully integrated in EFSA assessments, no further considerations regarding nutrients are included in this Opinion.

Another special case relates to hormesis, which refers to a biphasic dose-response to an environmental agent characterized by a low dose stimulation or beneficial effect and a high dose inhibitory or toxic effect. (e.g. Calabrese and Baldwin, 2001). This beneficial effect could, for example, be due to an adaptive or over-compensatory response to a chemical stressor (Calabrese, 2005). Chemical risk assessment concerns food safety and not the evaluation of beneficial effects, therefore hormesis is not considered in detail in this Opinion.

2. Data and Methodologies

2.1 Data

In line with the ToRs, the main data sources are the Report (Beausoleil et al., 2016.) and the follow up probabilistic assessment (Chevillotte et al. 2017a,b). All studies fulfilling 5 or 6 checkpoints in the Report have been included in the assessment, as well as the probabilistic assessments for these datasets.

In addition, it was considered appropriate to conduct an additional search for recent scientific literature on the topic. The available resources did not allow performance of a new systematic review, thus a targeted literature search for gathering additional relevant peer-reviewed publications between 2017 and October 2019 was conducted in November 2019. The details of this search and main findings are summarised in Table 1. The references and citations of the retrieved articles were also searched and relevant studies retrieved and included in the search.

Table 1 Characteristics and results of the complementary literature search

Database	String	Complementary search	Results
Web of Science selecting the	TS=(monotonic OR nonmonotonic OR	The search was complemented with	• 225 articles retrieved

[5] http://www.efsa.europa.eu/sites/default/files/consultation/consultation/Draft-statement-on-HBGV-for-PC.pdf

following indexes: SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC.	non-monotonic) AND TS=(toxic* AND dose	the analysis of the references and citations of the retrieved publications	• 19 additional experimental studies selected as final result after the screening

The 19 additional experimental studies were grouped according to the relevance of the tested chemical for EFSA. Six studies on BPA and six studies on phthalates, were considered relevant for this assessment. The other seven studies had been conducted with mixtures and with chemicals outside the EFSA remit, and were not further considered for this assessment.

In the Report (Beausoleil et al. 2016), BPA and two phthalates, di(2-ethylhexyl) phthalate (DEHP) and di-n-butyl phthalate (DBP), are the substances under EFSA remit with the highest number of *in vivo* datasets reporting potential NMDR (35 for BPA, 30 for DEHP and 5 for DBP). However, only for one of these datasets, aromatase activity in rats exposed to DEHP (Andrade et al., 2006), the six checkpoints were met. Considering the concordance between the Report and the complementary search, additional assessments regarding NMDR claims for BPA and phthalates have been performed and included as annexes to this scientific opinion.

Regarding previous EFSA risk assessments, tropane alkaloids were identified from an Opinion of the EFSA Scientific Panel on Contaminants in the Food Chain (EFSA CONTAM, 2013), as an example of a biologically relevant NMDR and included in this assessment. It should be noted that relevant publications will inevitably have been missed, as the term NMDR is often not used to describe these types of dose-response curves.

2.2 Methodologies

The methodology used by the authors in the Report (Beausoleil et al., 2016) and in the probabilistic assessment (Chevillotte et al., 2017a,b) has been briefly summarized in the Introduction (see Systematic review and probabilistic assessment subsections). To compare the consistency between the two methods that have been developed to assess NMDR (Beausoleil et al., 2016 and Chevillotte et al. 2017a,b), the results from the visual/statistical analysis of datasets judged to show potential NMDR (≥5 checkpoints) by the Report were compared with the probabilistic analysis conducted according to the methodology proposed by Chevillotte et al. (2017a,b). The probabilistic assessment, according to the Chevillotte et al. (2017a,b) methodology, has been also applied to additional datasets selected from EFSA assessments and publications retrieved in the complementary literature search.

The biological relevance of potential NMDRs identified was assessed by expert judgment, analysing each selected publication. The systematic approach developed considered three key elements: a) the role of the measured effect in the Adverse Outcome Pathway (AOP), distinguishing between early event, intermediate events and apical effects; b) the biological plausibility for a non-monotonic dose response, considering the measured effect and information on the mechanistic pathway when available; and c) the role in adversity for the observed NMDR, considering the principles for selecting the Reference Points (RP) for establishing Health-Based Guidance Values in EFSA guidance documents and its

implementation (information on the endpoints selected as RP in EFSA assessment is available from OpenFoodTox and additional details for pesticides were provided by the relevant unit).

The draft Scientific Opinion will be issued for public consultation, and comments will be assessed by the Working Group during the finalisation of the scientific opinion.

3. Assessment

The assessment is divided in two sections. Section 3.1 covers the *in vivo* studies included in the Report and containing datasets that fulfil five or six of the checkpoints in the visual/statistics analysis. Section 3.2 discusses other studies identified as potentially relevant from other EFSA activities but not covered in the Report, and summarises the evaluations done for BPA and phthalates, which are detailed in Annex A and B, respectively. One dataset from the Report meeting the six checkpoints addressing DEHP effects on aromatase inhibition in rats (Andrade et al., 2006) is included in the phthalates assessment (Annex B) instead of in Section 3.1.

3.1 *In vivo* studies with datasets fulfilling five or six checkpoints

This section briefly describes examples of datasets from the Report showing signs of non-monotonicity, in order to highlight possible differences in mode of action that may account for the observed non-monotonicity. The discussion is not meant to give a complete or thorough review but rather to set the stage for the examples summarized in Tables 2 and 3.

The Report covers a variety of studies addressing different measured effects. In some cases, the observed NMDR was considered to be caused by a well-known biological phenomenon, with intrinsically high biological plausibility for non-monotonicity. These observations are included in Table 2 and the presented data reflects two different processes that may explain the underlying NMDR. The first set of data covers responses considered as protective or of beneficial nature; such as the protective effect of resveratrol against induced gastric ulcer (Dey et al., 2009), the use of rosmarinic acid as an anxiolytic/antidepressant (Takeda et al., 2002), or of tanshinone IIA as an anticonvulsant (Buenafe et al., 2013). This form of non-monotonicity can be explained by two different mechanisms, the protective or beneficial effects observed at the lower doses are reduced and disappear at higher doses following the induction of toxicity. The second group covers those measuring motor stimulation and social investigation in experimental animals. Caffeine (Halldner et al., 2004; Marin et al., 2011; Zhang et al., 2011) and ethanol, including its metabolite acetaldehyde (Escarabajal and Aragon, 2002; Correa et al., 2003; Varlinskaya and Spear, 2009), provoked behavioural/locomotor stimulation, with NMDRs related to inhibition of the stimulation or even depression at higher doses. This is considered biologically plausible, as stimulation is expected to peak at a certain level and then may be affected by other biological responses (see Ferré et al., 2018, for a review on the modes of action for the induction and inhibition of locomotor activity by caffeine). The capacity of nicotine to both activate and desensitize/inactivate nicotinic acetylcholine receptors (nAChRs) is another well characterised phenomenon (Picciotto et al., 2008). The effects of metabolites, may also play a role at higher doses explaining the observed NMDR as suggested in the study by Escarabajal and Aragon (2002). The study by Bai and Zhu (2010), measuring the stimulatory effect of two bioflavonoids on

COX-mediated formation of PGE2 has been also included in this list, as it is linked to the stimulation of an intermediate event and the aim is to assess possible beneficial effects.

The biological plausibility of NMDR in the area of developmental neurotoxicity (DNT) has been addressed in the NAFTA DNT Guidance (Moser et al., 2016). Biologically plausible observations are confirmed for assessment of motor activity and auditory startle. The excitation followed by sedation produced by ethanol is a classic example (Moser et al., 2016). Neural systems reflect interplay of both inhibitory and excitatory actions, and the relative influence of these factors may impact a dose response. These may be observed as U-shaped or inverted U-shaped curves (Moser et al., 2016).

Table 3 presents the assessment of the other datasets in the Report meeting five or six checkpoints, covering a variety of different chemicals and measured effects where the underlying biology was considered less clear compared to those presented in Table 2. Each dataset with possible NMDR is analysed regarding biological plausibility and role in adversity.

346 **Table 2**. Studies fulfilling five or six "checkpoints" in the report by Beausoleil et al. (2016) for which a well-defined biological explanation for NMDR could
347 be identified.

Publication, chemical, and measured effects	**Dose range, # of dose-groups (N) excluding controls**	**1. Presence/ shape of NMDR[η]** (checkpoint not fulfilled*)	**2. Nature of measured effect**	**3. Biol plausibility[‡]**	**4. Role in adversity[‡]**	**5. Probability of NMDR (P_{NMDR} %) as described by Chevillotte et al. (2017a,b)**	**Comments**
Dey et al., 2009. Impact of **resveratrol** on indomethacin-induced gastric ulcer in mice 1. Ulcer index 2. Myeloperoxidase (MPO) activity	0.5-10 mg/kg p.o. starting the first dose 6 h after indomethacin administration N=6	1. Yes U (none) 2. Yes U (none)	1. Apical (beneficial) effect 2. Intermediate	1 Yes 2 Yes	1. Decrease in protective effect observed at higher doses 2. Marker of neutrophil aggregation at the site of inflammation, associated to ulcerated conditions and reduced with the healing process	1. P_{NMDR} 99.98 (result after 3 days) 2. P_{NMDR} 99.89 (results after 2 days)	Ulcer index and MPO were measured at different time points, probability values are reported for one time point The MOA was investigated The lower dose of resveratrol augmented eNOS expression without altering COX-1 expression, but, at a higher dose resveratrol predominantly suppressed COX-1 expression, which significantly reduced both PGE2 synthesis and angiogenesis.
Takeda et al., 2002. Impact of **rosmarinic acid** on freezing behaviour of mice exposed to a conditioned fear stress (inescapable electric foot shocks) 1. Duration of immobility	0.25-4 mg/kg i.p single dose N=5	1. Yes U (CP-5)	1. Apical effect	1 Yes	1Unclear, is an alteration of the natural response to stress Spontaneous motor activity was not affected.	1. P_{NMDR} 78.35 (result after 3 days)	Conditioned fear stress induced freezing behaviour is the period of crouching and complete immobility of rodents previously exposed to aversive stimuli such as inescapable foot-shocks. This is a stress model reflecting emotional abnormality including anxiety and/or depressive state and is attenuated by anxiolytics and antidepressants
Buenafe et al., 2013. Anticonvulsant activity of **Tanshinone IIA** in mice subjected to electrical stimulus through the corneas.	0.1-10 mg/kg i.v. N=5	1. Yes ∩ (CP-5)	1. Apical effect	1Yes	1. Decrease in protective effect observed at high doses	Not analysed	No effects at 0.1, 5 and 10 mg/kg, same effect at 0.5 and 1 mg/kg iv Biphasic/hormetic dose responses have indeed been previously reported in chemically diverse pro- and anticonvulsant agents with different modes of action

1. Number of mice protected							
Halldner et al., 2004. Impact of **caffeine** on locomotor activity in mice 1. Horizontal activity (number of counts indicating movements to adjancent cells)	3.75-100 mg/kg ip N=5	1. Yes ∩ But increase observed at all doses except the highest (CP-3)	1. Apical	1. Yes	1. Stimulation/Unclear role in adversity	1. P_{NMDR} 99.36 (result after 3 days)	Dose basing not optimal for assessing NMDR, Blockade of the adenosine A(2A) receptor (A2AR) is necessary for the stimulatory effect of low doses. The inhibitory effect of high doses is due neither to blockade of the A1R, nor of the A2AR, and an effect independent of these adenosine receptors is likely
Marin et al., 2011. Impact of **caffeine** on locomotor activity in rats 1. Horizontal activity adults (number of counts indicating movements to adjacent cells) 2. Horizontal activity adolescents (number of counts indicating movements to adjacent cells)	3-120 mg/kg ip N=5	1. Yes ∩ (CP-3) 2. Yes ∩ (CP-3)	1. Apical 2. Apical	1. Yes 2. Yes	1 Stimulation/Unclear role in adversity 2. Stimulation/Unclear role in adversity	1. P_{NMDR} 99.41 2. P_{NMDR} 88.27	Antagonism of A2A receptors is clearly related to stimulant properties of caffeine. High caffeine doses also act on less specific cellular targets other than adenosine antagonism. These mechanisms include the inhibition of phosphodiesterase enzyme, blockade of GABAA receptors or mobilization of calcium from intracellular stores (Fisone et al., 2004)
Zhang et al., 2011. Impact of **caffeine** on locomotor activity in mice 1. Horizontal activity (travel distance)	1-100 mg/kg ip N=5	1. Yes ∩ (CP-3) 2. Yes ∩ (N/A)	1. Apical 2. Apical	1. Yes 2. Yes	1. Stimulation/Unclear role in adversity 2. Stimulation/Unclear role in adversity	1. P_{NMDR} 99.82 (result after 3 days) 2. Not analysed	Theophylline exhibited a similar but smaller decrease at higher doses.

2. Distance ratios in central and periferal regions							
Correa et al., 2003. Impact of **ethanol and its metabolites** on locomotor activity in rats 1. **Ethanol** induced horizontal activity(number of counts indicating movements to adjacent cells) 2. **Acetaldehyde** induced horizontal activity (number of counts indicating movements to adjacent cells)	1. Ethanol 16-258 microg intracranial injection N=5 2. Acetaldehyde 15-247 microg intracranial injection N=5 3. Acetate 21-168 microg intracranial injection N=5	1. Yes ∩ (CP-3) 2. Yes ∩ (CP-3)	1. Apical 2. Apical	1. Yes 2. Yes	1. Stimulation/Unclear role in adversity 2. Stimulation/Unclear role in adversity	1. P_{NMDR} 88.43 2. P_{NMDR} 79.33	Acetate induced monotonic inhibition in horizontal activity (number of counts indicating movements to adjacent cells) Results suggest that some of the motor suppression or sedation produced by ethanol at high doses could be related to the metabolite acetate
Escarabajal and Aragon, 2002. Impact of **ethanol** on motor activity in mice 1. Horizontal activity (number of counts indicating movements to adjacent cells)	0.8-4 g/kg ip injection N=5	1. Yes ∩ (CP-5)	1. Apical	1. Yes	1. Stimulation	1. P_{NMDR} 99.79	Cyanamide, a catalase and ALDH inhibitor suppressed the NMDR of ethanol. The antidote 4-methylpyrazole (4-MP), an alcohol dehydrogenase (ADH) inhibitor, enhanced the NMDR of ethanol
Varlinskaya and Spear, 2009. Impact of **ethanol** on motor activity in mice	0.25-1.25 g/kg sc injection N=5	1. Yes ∩ but only at 1 dose (CP-3) 2. Yes ∩ (CP-3)	1. Apical 2. Apical	1. Yes 2. Yes	1. Stimulation 2. Stimulation	1. P_{NMDR} 97.37 2. P_{NMDR} 95.06	To note that locomotor activity was not affected by ethanol in this study The nonselective opioid antagonist naloxone and the selective μ-opioid antagonist CTOP blocked the stimulatory effects of ethanol

1. Behavior as social investigation 2. Behavior as play fighting							on play fighting but not on social investigation.
Bai and Zhu, 2010. role of two bioflavonoids as co-substrates for cyclooxigenases (COX) in rats **1.** Impact of **myricetin** on PGE2 levels plasma **2.** Impact of q**uercetin** on PGE2 levels plasma	0.05-5 mg/kg bw day N=5	1. Yes ∩ (CP-5) 2. Yes ∩ (none)	1 & 2. Intermediate	1 &2. Unclear as not consistent with previous literature (see comment)	1&2. Stimulatory effect on COX-mediated formation of PGE2	1. P_{NMDR} 92.24 2. P_{NMDR} 99.89	Both stimulation and inhibition of COX-mediated formation of PGE2 may trigger other responses. Previous literature suggests inhibitory effect of bioflavonoids on COX activity

*CP = checkpoint as defined in the Report:
CP-3. Can the apparent NMDR be explained by one single potential outlying dose group?
CP-5. Is the steepness of the dose-response curve outside the range of biologically plausible/realistic dose-response shapes?
∏ The symbol U indicates a NMDR with U (or J) shape, the symbol ∩ indicates a NMDR with inverted U (or J) shape
‡ Only addressed when a possible NMDR is confirmed under 1. Presence/ shape of NMDR

360 **Table 3**. Other studies fulfilling five or six "checkpoints" in the report by Beausoleil et al. (2016) for which well-defined biological explanation for NMDR
361 were subject to some uncertainty.

Publication, chemical and measured effect	**Dose range, # of dose-groups (N) excluding controls**	**1. Presence/ shape of NMDR** (checkpoint not fulfilled*)	**2. Nature of measured effect**	**3. Biol plaus‡**	**4. role in adversity‡**	**5. Probability of NMDR (%) as described by Chevillotte et al. (2017a,b)**	**Comments**
Puatanochochai et al., 2006. Impact of **alpha HCH** on hepatic markers in rats pre-induced with diethylnitrosamine 1. Proliferation of GST-P positive hepatic foci 2. Total CYP450 content in liver 3. Proliferating-Cell-Nuclear-Antigen (PCNA) 4. 2α-testosterone hydroxylase activity in liver 5. 8OHdG formation in liver 6. NADPH-P450 reductase activity in liver 7. 16α-testosterone hydroxylase activity in liver	0.01-500 mg/kg diet 10wk (0.001-50 mg/kg bw) N=7 All rats had received 100 mg/kg bw ip diethylnitrosamine weekly 3 times before starting alpha HCH exposure	1. No (CP-3) 2. Yes U (CP-3) 3. Yes U (none) 4. Yes ∩ (none) 5. Yes U (none) 6. Yes U (none) 7. Yes ∩ (CP-5)	1. Intermediate 2. Early event 3. Early event 4. Early event 5. Intermediate 6. Early event 7. Early event	?	1. Yes 2. No 3. Yes, together with cell proliferation 4. Unclear 5. Decrease is protective, increase is adverse 6. Unclear 7. Unclear	1. P_{NMDR} 77.0 (U) 2. P_{NMDR} 92.0 3. P_{NMDR} 99.14 4. P_{NMDR} 99.96 5. P_{NMDR} 89.37 (U) 6. P_{NMDR} 97.39 (U) 7. P_{NMDR} 79.5 (∩)	Could be related to combined effect of the two substances Four checkpoints met for CYP2C11 mRNA expression in liver 4 and 7. Monotonic increases for other testosterone hydroxylase activities
Zhang et al., 2012. Acute effects of **methylmercury** ip on rats	2-10 mg/kg bw ip, 1x N=6	1. Yes, ∩ but toxicity could explain the decrease in	1. Early event	1. Yes	1. Unclear	1. P_{NMDR} 72.0	Not relevant for the much lower human exposure. Furthermore, acute ip application

1. Protein expression in cerebral cortex as marker for stress response		protein expression at doses >6 mg/kg (none)					
Shutoh et al., 2009. Effects of **DDT** on juvenile rats 1. DNA methylation, and indicators of oxidative stress (lipid peroxidation; LPO) in cerebrum	0.06-60 mg/kg bw 4wk Gavage, N=6	1. Yes U for LPO, other changes not convincing (CP-3)	1. Early event	1. Yes	No. Homeostatic response to a xenobiotic	1. P_{NMDR} 87.85	
DDT Sukata et al., 2002. Effects of **DDT** on rats. 1. Proliferation of GST-P positive hepatic foci 2. mRNA IL-1 receptor type 1 (Fig3)	0.5-500 mg/kg diet 16 wk (0.05-20 mg/kg bw) N=8	1. No (CP-3) 2. Yes, trend, not stat. Sign. ∩ (CP-3)	1. Intermediate 2. ?		Rather an indication of induction of anti-stress responses at low doses	1. P_{NMDR} 77.35 (NMDR U) (2 cells) 2. P_{NMDR} 83.86	1. GST-P positive foci of different size classes were analysed 2. Similar result for other mRNA
Yuanqing et al., 2013 Effects of **acetonitrile** on mice. 1. AChE brain	0.156-20 mg/kg N=8 i.p. adm	1. Yes U (CP-3)	1. Intermediate effect, but has been used as RP	1. ?	Inhibition has been used as RP for adversity	1. P_{NMDR} 100	Four checkpoints for AChE blood with ∩
Wildemann et al., 2015 Effects of **lead acetate** on rats 1. Body weight gain 2. Pulse pressure	0.004-45 mg/kg bw/d N=8 Drinking water	1. Yes, ∩ (CP-5) 2. Yes, U	1. Apical 2. Intermediate	?	1. Yes, body weight gain was 113 g control vs up to 224g treated 2-7 Yes	1. P_{NMDR} 92.38 2. P_{NMDR} 76.64	All the hemodynamic effects are linked. Other possible non-monotonic responses but with less than 5 checkpoints observed for Systolic blood pressure Stroke volume Cardiac output

<u>Zorrilla et al., 2009</u> Effects of **triclosan** on juvenile rats 1. Triiodothyronine (T3) serum	3-300 mg/kg/day N=5 Gavage	1. Yes, ∩ (CP-3)	1. Intermediate	1. ? 3. No	1. Yes, reduction in T levels during critical windows is linked to reproductive effects	1. low for NMDR (56% for **MDR**)	1. Due to one dose group, but very high reduction. Large variability among treatments 3. The main effect is for T4 and is clearly monotonic

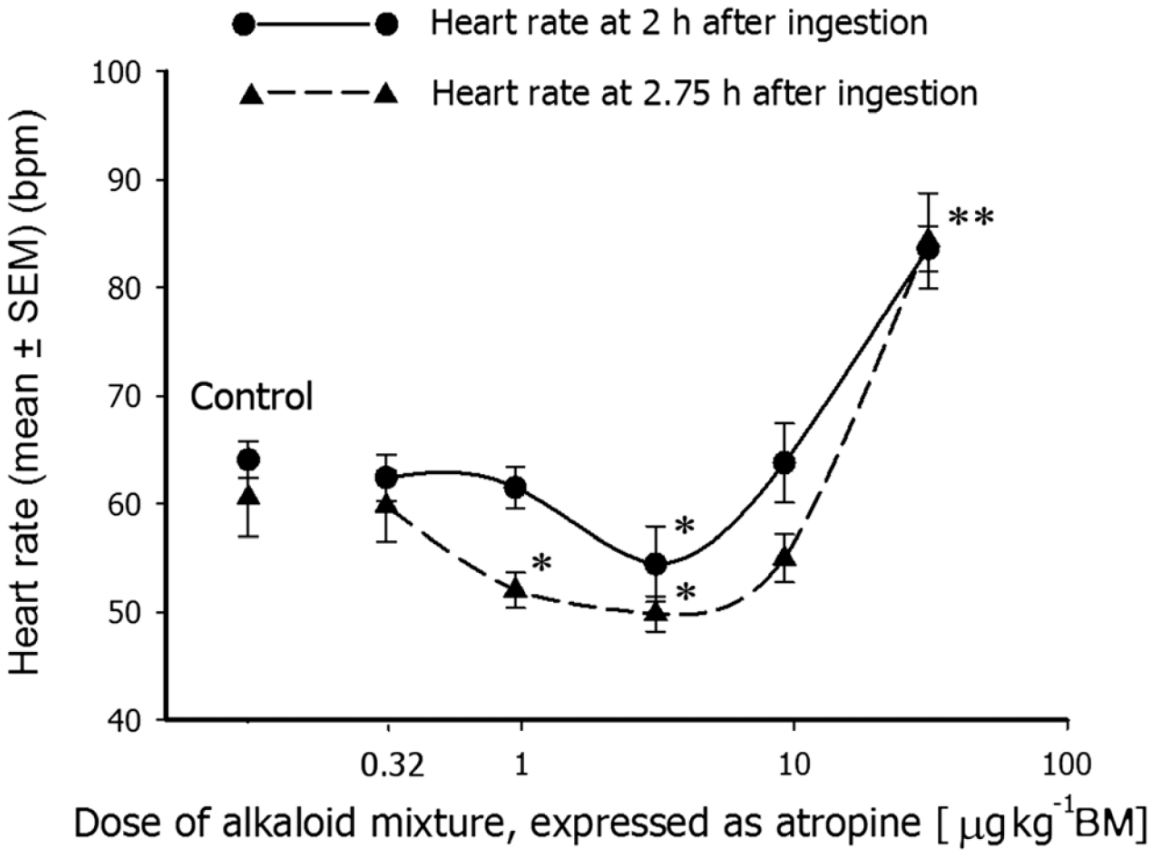
*CP = checkpoint as defined in the Report:
CP-3. Can the apparent NMDR be explained by one single potential outlying dose group?
CP-5. Is the steepness of the dose-response curve outside the range of biologically plausible/realistic dose-response shapes?
∏ The symbol U indicates a NMDR with U (or J) shape, the symbol ∩ indicates a NMDR with inverted U (or J) shape
‡ Only addressed when a possible NMDR is confirmed under 1. Presence/ shape of NMDR

Consistency between the different approaches is observed throughout Table 2, which describes cases with a well-defined biological explanation for the NMDR. The probability for NMDR according to the methodology described by Cheviollotte et al. (2017a,b) was higher than 78% in all cases, and the NMDR confirmed by the expert judgement.

Two checkpoints, CP-3 (Can the apparent NMDR be explained by one single potential outlying dose group?) and CP-5 (Is the steepness of the dose-response curve outside the range of biologically plausible/realistic dose-response shapes?) were not meet for some datasets with high likelihood for NMDR in the probabilistic assessment. Other discrepancies between the two methodologies were observed in some cases, confirming that each method provides information on different elements. In two cases, the expert judgement concluded that there were no indications for NMDR, despite the dataset fulfilled five checkpoints and the likelihood in the probabilistic analysis was higher that 75%. The biological plausibility was clear for all datasets reported in Table 2, but remained doubtful for the majority of datasets reported in Table 3.

3.2 Other studies

Tropane alkaloids were identified from an Opinion of the EFSA Scientific Panel on Contaminants in the Food Chain (EFSA CONTAM, 2013), as an example of a biologically relevant NMDR. These alkaloids are present in various plant species that can contaminate food-producing plants. The main tropane alkaloids, hyoscyamine and scopolamine, exhibit anticholinergic activity, due to competitive inhibition of acetylcholine binding to muscarinic receptors. This results in a number of pharmacological effects including salivary secretion, pupil dilation and heart rate changes. The effect on heart rate is biphasic (see Figure 2), with a decrease at lower doses and increase at higher doses. The mode of action has been previously discussed (Pitschner and Wellstein, 1988; Wellstein and Pitschner, 1988; Pitschner et al., 1994). Both of these effects were covered in the risk assessment by using the NOAEL for decreased heart rate as the reference point for establishing an Acute Reference Dose.



394

395 **Figure 2**. Dose-response curve for heart rate versus the dose of the atropine/scopolamine
396 mixture, expressed as atropine (*p<0.005, **p<0.001). Reproduced with permission from
397 Perharic et al., 2013 (DOI 10.1002/jat.2797)

398

399 Results of additional probabilistic assessments for Perharic et al. (2013) conducted according
400 to the methodology proposed by Chevillotte et al. (2017a,b) confirm the NMDR with associated
401 probabilities for a U shaped dose-response of 66.1% and 86.7% at 2 and 2.75h respectively.
402

403 The external Report identified four studies on **BPA** where possible NMDR had been examined
404 or claimed. That is, studies were not necessarily picked up because they provided convincing
405 evidence of NMDR but rather because the word came up in the publication. As an example, a
406 study by Tyl et al (2002), was identified as the study was designed to examine possible NMDR
407 for developmental effects of BPA. Although the authors concluded in their publication that no
408 indication of NMDR was present in their results, liver weight in the second generation (F2)
409 was still evaluated by the Report. The results being in line with those of the authors that
410 presence of NMDR was unclear (only three checkpoints were fulfilled). The SC evaluation
411 reached the same conclusion (See Annex I). Other studies on BPA identified by the Report
412 included possible NMDR for extracellular kinase signalling in cerebellar cortex (pERK-
413 IRCellAtP10) (Zsarnovszky et al., 2005), semen quality (Kendig et al., 2012); and gonadal and
414 renal fat pads (Angle et al., 2013). Only four checkpoints were fulfilled for each of these
415 studies. For risk assessment the relevance of an effect on extracellular kinase signalling in
416 cerebellar cortex, in the absence of other functional measures, remains unclear. For effects

on semen quality, the possible NMDR observed in the study by Kendig et al. (2012) was an inverted U-shaped dose-response meaning, if anything, improved semen quality in the middle of the dose range which then went back to control level at higher doses. The study on renal and gonadal fat pads showed some suggestion of higher weight at low doses following prenatal exposures.

For risk assessment the effects on semen quality, renal or gonadal fat pads or other measures of adiposity would be of relevance. To address these findings for BPA further, a more targeted search for studies on BPA showing possible NMDR for these outcomes was conducted. Publications from the CLARITY-BPA programme (Consortium Linking Academic and Regulatory Insights on BPA Toxicity) were evaluated as well. One publication reported no effects on sperm quality (Camacho et al., 2019), another on more detailed sperm endpoints reported an inverted U-shaped dose-response for sperm DNA methylation with no indication of adversity for other semen parameters (Dere et al., 2018). A previous study had reported a possible U-shaped NMDR for sperm count (Hass et al., 2016), but the effect size observed in that study was modest. Overall, the presence of NMDR for sperm quality seems unlikely.

There were some indications of NMDR for gonadal fat pads following prenatal exposures to BPA (Taylor et al., 2018). These results are in line with those reported in Angle et al., (2013) but with only three dose groups, a proper evaluation of NMDR is not possible. A recent paper by Uchtmann et al. (2020) from the Clarity project concluded that, after exclusion of few animals (considered as outliers), there was an inverted U-shaped NMDR in body weight in offspring exposed to BPA in utero at postnatal day 1. No such results were observed at later ages. Our own statistical evaluation could not confirm that conclusion. Overall, the possible NMDR on measures of body composition seem unstable due to high variability across dose groups and modest effect size.

Finally, a few other reports from the Clarity project have suggested some indications of NMDR. The outcomes assessed, including different measures of fetal urogenital sinus (Uchtmann et al., 2020), mammary gland response (Montevil et al., 2020), percent basophils and modest changes in changes in % basophil and serum bile acid concentrations (Badding et al., 2019). Overall, due to the modest effect sizes observed without clear changes in other related biomarkers, the relevance of these findings for risk assessment is unclear and these findings need to be replicated for further evaluation.

An additional probabilistic assessment for NMDR (see Appendix A for details) was conducted for several datasets extracted from Uchtmann et al. (2020) according to the methodology proposed by Chevillotte et al. (2017a,b). For body weight the probability for NMDR is 58.8%, while for colliculus angle (litter) at PND1, and urogenital sinus epithelium thickness (midway section), monotonic dose-responses have higher probabilities than NMDRs.

The Report identified, using the statistical/visual approach, a NMDR for **DEHP** on aromatase activity, and there are a number of publications claiming NMDR for phthalates and DEHP in particular. The assessment included in Annex II revealed that the focus should be on testosterone levels and DEHP exposure covering development and pubertal exposure windows. There is a connection with the NMDR observed in the Report for aromatase as this enzyme is involved in testosterone metabolism.

There is experimental evidence supporting that the NMDR observed for this intermediate effect could be related to the disturbance of the hypothalamic–pituitary–gonadal axis (HPG)

feedback mechanism. A possible mechanistic interpretation is the combination of two different phenomena: first, the phthalate induced reduction in testosterone production capacity by Leydig cells, and second the compensatory Leydig cells hyperplasia triggered by the feedback mechanism. This could result in a NMDR for testosterone levels. A reduction of testosterone levels during a critical period could result in adverse reproductive effects. The net increase in testosterone levels will not result in these reproductive effects, but may be connected to different adverse outcomes. The critical period for adverse outcomes and the critical exposure window for an increase in testosterone, may be different from those related to the decrease in testosterone levels. The possible hypothesis to be explored is that phthalates may have a NMDR for the intermediate event, testosterone level, linked to different routes of adverse outcomes. High phthalate exposure produces anti-androgenic effects linked to testosterone decrease and the associated reproductive adverse outcomes (mostly malformations). Low phthalate exposure levels may be linked to testosterone increases (i.e. through the overstimulation of the compensatory feedback mechanism) and may be associated with different adverse outcomes, e.g. signalling to masculinization in females (i.e. hirsutisms) and neurodevelopmental effects associated with testosterone increase.

In conclusion from the analysis in Annex II is seems clear that the observed NMDR is caused by two different modes of action. For risk assessment the effects occurring in the lower dose range will be the critical one, being protective for the effects occurring at higher doses; however, such an assessment is outside the scope of this mandate.

3.3 Impact of the observed NMDR on the risk assessment process

Risk assessment of chemicals in food comprises the four steps of hazard identification, hazard characterisation (including dose–response assessment), exposure assessment and risk characterisation. NMDR could impact the risk assessment process at the hazard characterisation step, i.e. the identification of a reference point (RP) during the dose–response assessment. In principle, NMDR may occur at any region of the dose-response curve. Non-monotonicity occurring at the high-dose end of the dose-response curve does not impact the current hazard characterisation as the RP to establish a HBGV or calculating a MOE is the lowest dose where adverse effects can be observed and thus the RP would not change because of effects occurring at high doses. Furthermore, high-dose-effects are often caused by saturation effects or by overt toxicity impacting on the endpoint under consideration. NMDR may also be explained by different modes of action (MOA) operating at different dose-levels (see section 3.1.1.). This includes the induction of additional MOAs at high doses, e.g. via the production of toxic metabolites when detoxication pathways of the compound under consideration are overwhelmed. This will also not impact the hazard characterisation step.

Non-monotonicity occurring at the low-dose end of the dose-response curve could impact the current hazard characterisation particularly when an apical endpoint is affected. However, presently all the identified NMDR for non-nutrients observed *in vivo*[6] concern either early or intermediate events in the toxicity pathways, not leading to non-monotonicity of the related apical endpoints usually used for identifying a RP. During the evaluation of these NMDR, it is

[6] NMDR may also be observed in *in vitro* studies. However, *in vitro* studies are often mechanistic studies and not currently used as a basis for establishing HBGV. *In vitro* studies are not further considered here in line with the Terms of Reference provided in the mandate; nevertheless, as indicated below, NAMs including *in vitro* studies may provide the mechanistic information required for understanding the pathway to adversity for NMDR.

necessary to consider the biological relevance of the early or intermediate effects and the potential consequences of the effect (i.e. the potential for leading to adversity). When early or intermediate events are considered being adaptive physiological (or homeostatic) responses, no adverse effects are to be expected and thus would also not impact the hazard characterisation step. Some early or intermediate effects may be even beneficial (e.g. induction of DNA repair enzymes may lead to an improved repair of endogenous DNA lesions). Only when those early or intermediate events trigger further events leading to adverse effects, i.e. being biomarker of adverse effects, these should be taken into account in the hazard characterisation as it is done for monotonic dose-responses (e.g. β_2-microglobulin excretion in the kidney induced by cadmium). As another example may serve receptor mediated effects: it is well established that compounds interacting with cellular receptors may lead to bi-phasic effects. While lower doses stimulate the receptor, higher doses may block it, leading to opposite effects and may be considered as NMDR. Such effects are common in pharmacology and should be addressed in the hazard characterisation by identifying a pharmacological RP to establish a pharmacological HBGV, if this RP represents the most sensitive effect.

Overall, in evaluating a substance for which information on NMDR relations for one or more outcomes is obtained, the current risk assessment approach based on evaluating adverse outcomes seen in standard animal tests (as well as other observations) remains valid. With this in mind, the process recommended to be followed in cases of non-monotonicity is the following:

- Consider at which end of the dose-response curve non-monotonicity is observed:

  - If at the upper end of the dose-response curve, follow the current approach for determining a RP and establishing an HBGV.

  - If at the lower end, further considerations need to be taken into account as follows:
    - Is the effect observed an apical effect and is supported by further experimental work? If no, further investigations are needed.
    - If the observed effect is an early or intermediate effect, consider:
      - What is the evidence for the effect observed (*in vitro/in vivo*? Other?).
      - What is the biological relevance of the effects observed? Can a (quantitative) relation between these effects and an adverse outcome (i.e., apical effect) be established? Ideally: Could a mechanistic sequence (AOP) be partially or fully established? If yes, specific considerations need to be applied and a diversion from the current methodologies for RA as described in EHC 240 (IPCS, 2009) or FOSIE (Barlow et al., 2002) may be needed.
      - If information is lacking on whether an observed effect can lead to an adverse outcome, additional testing may be needed. Here New Approach Methodologies (NAMs) would be of relevance given the need for identifying a mechanistic sequence of events.

In cases where biological considerations or previous results suggest that NMDR may be present, any further testing should assure that a sufficient number of doses are tested at the lower end of the dose-response curve with an adequate dose-spacing to enable identifying potential NMDR. If such design issues are not properly considered, the possible presence or non-presence of NMDR cannot be addressed. Inclusion of sufficient number of dose groups

would also benefit the application of the benchmark-dose (BMD) approach. Furthermore, mechanistic data would inform whether or not early/intermediate effects show non-monotonicity.

4. Conclusions

Non-monotonic dose-response relations identified via the checkpoints approach and/or the probabilistic methodologies were reviewed, and their biological relevance assessed. The information compiled in the Report by Beausoleil et al. (2016) and the publications from Chevillotte et al. (2017a,b) were complemented with targeted literature searches and previous EFSA examples. Overall, it was concluded that:

- In assessing dose-response relationships for non-monotonicity, the checkpoint approach may yield a different result than those obtained through probabilistic (statistical) methodology;
- There is currently no gold standard for the statistical assessment of NMDR for chemical risk assessment. Therefore, using different statistical approaches may result in diverging conclusions when used individually;
- Apparent NMDR have been observed in a number of studies with different chemicals using three approaches (checkpoints, probabilistic assessment and expert judgment);
- Apparent NMDR are observed for early (molecular) or intermediate events, but also for some apical effects relevant for the risk assessment;
- If an NMDR is observed for an apical effect, the understanding of the underlying mechanism(s) is necessary to assess its biological plausibility and to consider the consequences for the risk assessment process;
- An NMDR in an apical effect may result from two or more modes of action, each with a monotonic dose response. If the effect observed at lower doses is considered adverse, this effect would be selected to identify the RP for risk assessment. A special case is encountered in the case of nutrients with two independent dose-response curves observed: one for deficiency and another for toxicity; the adverse effects on both sides are generally different;
- If an NMDR is observed for an early or intermediate event, the potential for propagating towards an apical effect needs to be demonstrated and checked for its biological relevance as above. It should be noted that molecular or intermediate events leading to effects in opposite directions may be linked to different adverse effects at apical level, each occurring at different exposure ranges and not showing an NMDR.

Taking into account the conclusions above, and in order to provide a way forward, a process to be followed for addressing NMDR in the risk assessment is outlined in chapter 3.3. This approach is recommended for application in cases of apparent non-monotonicity.

The approach was applied to two case studies: Bisphenol A (BPA) and Phthalates. No indications of NMDR have been detected for BPA, while for the phthalate DEHP, indications for a biologically plausible NMDR were observed for an intermediate effect, testosterone levels, possibly linked to the feedback control mechanism. The impact of this NMDR on the risk assessment of DEHP should be further investigated.

References

Andrade AJM, Grande SW, Talsness CE, Grote K and Chahoud I, 2006. A dose-response study following in utero and lactational exposure to di-(2-ethylhexyl)-phthalate (DEHP): Non-monotonic dose-response and low dose effects on rat brain aromatase activity. Toxicology, 227:185-192. doi: 10.1016/j.tox.2006.07.022

Angle BM, Do RP, Ponzi D, Stahlhut RW, Drury BE, Nagel SC, Welshons WV, Besch-Williford CL, Palanza P, Parmigiani S, vom Saal FS and Taylor JA, 2013. Metabolic disruption in male mice due to fetal exposure to low but not high doses of bisphenol A (BPA): evidence for effects on body weight, food intake, adipocytes, leptin, adiponectin, insulin and glucose regulation. Reprod Toxicol, 42:256-268. doi: 10.1016/j.reprotox.2013.07.017

Badding MA, Barraj L, Williams AL, Scrafford C and Reiss R, 2019. CLARITY-BPA Core Study: Analysis for non-monotonic dose-responses and biological relevance. Food Chem Toxicol, 131:110554. doi: 10.1016/j.fct.2019.06.001

Bai H-W and Zhu BT, 2010. Myricetin and quercetin are naturally occurring co-substrates of cyclooxygenases in vivo. Prostaglandins, leukotrienes, and essential fatty acids, 82:45-50. doi: 10.1016/j.plefa.2009.10.006

Barlow S, Dybing E, Edler L, Eisenbrand G, Kroes R and Brandt PA, 2002. Food Safety in Europe (FOSIE): Risk Assessment of Chemicals in Food and Diet. Food and Chemical Toxicology, 40

Beausoleil C, Beronius A, Bodin L, Bokkers BGH, Boon PE, Burger M, Cao Y, De Wit L, Fischer A, Hanberg A, Leander K, Litens-Karlsson S, Rousselle C, Slob W, Varret C, Wolterink G and Zilliacus J, 2016. Review of non-monotonic dose-responses of substances for human risk assessment. EFSA Supporting Publications, 13:1027E. doi: 10.2903/sp.efsa.2016.EN-1027

Buenafe OE, Orellana-Paucar A, Maes J, Huang H, Ying X, De Borggraeve W, Crawford AD, Luyten W, Esguerra CV and de Witte P, 2013. Tanshinone IIA exhibits anticonvulsant activity in zebrafish and mouse seizure models. ACS chemical neuroscience, 4:1479-1487. doi: 10.1021/cn400140e

Calabrese EJ and Baldwin LA, 2001. Hormesis: a generalizable and unifying hypothesis. Crit Rev Toxicol, 31:353-424. doi: 10.1080/20014091111730

Calabrese EJ and Mattson MP, 2017. How does hormesis impact biology, toxicology, and medicine? npj Aging and Mechanisms of Disease, 3:13. doi: 10.1038/s41514-017-0013-z

Calabrese EJ, 2005. Toxicological awakenings: the rebirth of hormesis as a central pillar of toxicology. Toxicology and Applied Pharmacology, 204:1-8. doi: https://doi.org/10.1016/j.taap.2004.11.015

Camacho L, Lewis SM, Vanlandingham MM, Olson GR, Davis KJ, Patton RE, Twaddle NC, Doerge DR, Churchwell MI, Bryant MS, McLellen FM, Woodling KA, Felton RP, Maisha MP, Juliar BE, Gamboa da Costa G and Delclos KB, 2019. A two-year toxicology study of bisphenol A (BPA) in Sprague-Dawley rats: CLARITY-BPA core study results. Food and Chemical Toxicology, 132:110728. doi: https://doi.org/10.1016/j.fct.2019.110728

Chevillotte G, Bernard A, Varret C, Ballet P, Bodin L and Roudot AC, 2017a. Probabilistic assessment method of the non-monotonic dose-responses-Part I: Methodological approach. Food Chem Toxicol, 106:376-385. doi: 10.1016/j.fct.2017.05.070

Chevillotte G, Bernard A, Varret C, Ballet P, Bodin L and Roudot AC, 2017b. Probabilistic assessment method of the non-monotonic dose-responses-Part II: Robustness assessment. Food Chem Toxicol, 110:214-228. doi: 10.1016/j.fct.2017.10.030

644 Conolly RB and Lutz WK, 2004. Nonmonotonic Dose-Response Relationships: Mechanistic
645 Basis, Kinetic Modeling, and Implications for Risk Assessment. Toxicological Sciences,
646 77:151-157. doi: 10.1093/toxsci/kfh007

647 Correa M, Arizzi MN, Betz A, Mingote S and Salamone JD, 2003. Open field locomotor effects
648 in rats after intraventricular injections of ethanol and the ethanol metabolites acetaldehyde
649 and acetate. Brain Res Bull, 62:197-202. doi: 10.1016/j.brainresbull.2003.09.013

650 Dere E, Anderson LM, Huse SM, Spade DJ, McDonnell-Clark E, Madnick SJ, Hall SJ, Camacho
651 L, Lewis SM, Vanlandingham MM and Boekelheide K, 2018. Effects of continuous bisphenol
652 A exposure from early gestation on 90 day old rat testes function and sperm molecular
653 profiles: A CLARITY-BPA consortium study. Toxicol Appl Pharmacol, 347:1-9. doi:
654 10.1016/j.taap.2018.03.021

655 Dey A, Guha P, Chattopadhyay S and Bandyopadhyay SK, 2009. Biphasic activity of resveratrol
656 on indomethacin-induced gastric ulcers. Biochemical and Biophysical Research
657 Communications, 381:90-95. doi: https://doi.org/10.1016/j.bbrc.2009.02.027

658 ECHA (European Chemicals Agency) and EFSA (European Food Safety Authority) with the
659 technical support of the Joint Research Centre (JRC), Andersson N, Arena M, Auteri D,
660 Barmaz S, Grignard E, Kienzler A, Lepper P, Lostia AM, Munn S, Parra Morte JM, Pellizzato
661 F, Tarazona J, Terron A and Van der Linden S, 2018. Guidance for the identification of
662 endocrine disruptors in the context of Regulations (EU) No 528/2012 and (EC) No
663 1107/2009. EFSA Journal 2018;16(6):5311, 135 pp.
664 https://doi.org/10.2903/j.efsa.2018.5311. ECHA-18-G-01-EN.

665 EFSA (European Food Safety Authority), 2010. Application of systematic review methodology
666 to food and feed safety assessments to support decision making. EFSA Journal 2010; 8(
667 6):1637. [90 pp.]. doi:10.2903/j.efsa.2010.1637

668 EFSA (European Food Safety Authority), 2012. EFSA's 17th Scientific Colloquium on low dose
669 response in toxicology and risk assessment. EFSA Supporting Publication 2012; 9(11):EN-
670 353, 64 pp. doi:10.2903/sp.efsa.2012.EN-353

671 EFSA CEP Panel (EFSA Panel on Food Contact Materials, Enzymes and Processing Aids), Silano
672 V, Barat Baviera JM, Bolognesi C, Chesson A, Cocconcelli PS, Crebelli R, Gott DM, Grob K,
673 Lampi E, Mortensen A, Rivière G, Steffensen I-L, Tlustos C, Van Loveren H, Vernis L, Zorn
674 H, Cravedi J-P, Fortes C, Tavares Poças MF, Waalkens-Berendsen I, Wölfle D, Arcella D,
675 Cascio C, Castoldi AF, Volk K and Castle L, 2019. Scientific Opinion on the update of the
676 risk assessment of di-butylphthalate (DBP), butyl-benzyl-phthalate (BBP), bis(2-
677 ethylhexyl)phthalate (DEHP), di-isononylphthalate (DINP) and di-isodecylphthalate (DIDP)
678 for use in food contact materials. EFSA Journal 2019;17(12):5838, 85 pp.
679 https://doi.org/10.2903/j.efsa.2019.5838

680 EFSA CONTAM Panel (EFSA Panel on Contaminants in the Food Chain), 2013. Scientific
681 Opinion on Tropane alkaloids in food and feed. EFSA Journal 2013; 11 (10):3386, 113 pp.
682 doi:10.2903/j.efsa.2013.3386

683 Escarabajal MD and Aragon CMG, 2002. The effect of cyanamide and 4-methylpyrazole on the
684 ethanol-induced locomotor activity in mice. Pharmacology, biochemistry, and behavior,
685 72:389-395. doi: 10.1016/s0091-3057(01)00762-6

686 Ferré S, Díaz-Ríos M, Salamone JD and Prediger RD, 2018. New Developments on the
687 Adenosine Mechanisms of the Central Effects of Caffeine and Their Implications for
688 Neuropsychiatric Disorders. Journal of caffeine and adenosine research, 8:121-130. doi:
689 10.1089/caff.2018.0017

690 Fisone G, Borgkvist A and Usiello A, 2004. Caffeine as a psychomotor stimulant: mechanism
691 of action. Cell Mol Life Sci, 61:857-872. doi: 10.1007/s00018-003-3269-3

692 Halldner L, Adén U, Dahlberg V, Johansson B, Ledent C and Fredholm BB, 2004. The adenosine
693 A1 receptor contributes to the stimulatory, but not the inhibitory effect of caffeine on
694 locomotion: a study in mice lacking adenosine A1 and/or A2A receptors.
695 Neuropharmacology, 46:1008-1017. doi: 10.1016/j.neuropharm.2004.01.014

696 Hass U, Christiansen S, Boberg J, Rasmussen MG, Mandrup K and Axelstad M, 2016. Low-
697 dose effect of developmental bisphenol A exposure on sperm count and behaviour in rats.
698 Andrology, 4:594-607. doi: 10.1111/andr.12176

699 Hill CE, Myers JP and Vandenberg LN, 2018. Nonmonotonic Dose-Response Curves Occur in
700 Dose Ranges That Are Relevant to Regulatory Decision-Making. Dose-Response, 16. doi:
701 10.1177/1559325818798282

702 IPCS (International Programme on Chemical Safety), 2002. Principles and methods for the
703 assessment of risk from essential trace elements. World Health Organization.
704 https://apps.who.int/iris/handle/10665/42416

705 IPCS (International Programme on Chemical Safety), 2009. Environmental Health Criteria 240,
706 Principles and Methods for the Risk Assessment of Chemicals in Food. World Health
707 Organization.

708 Kendig EL, Buesing DR, Christie SM, Cookman CJ, Gear RB, Hugo ER, Kasper SN, Kendziorski
709 JA, Ungi KR, Williams K and Belcher SM, 2012. Estrogen-like disruptive effects of dietary
710 exposure to bisphenol A or 17α-ethinyl estradiol in CD1 mice. Int J Toxicol, 31:537-550.
711 doi: 10.1177/1091581812463254

712 Lagarde F, Beausoleil C, Belcher SM, Belzunces LP, Emond C, Guerbet M and Rousselle C,
713 2015. Non-monotonic dose-response relationships and endocrine disruptors: a qualitative
714 method of assessment. Environ Health, 14:13. doi: 10.1186/1476-069x-14-13

715 Marin MT, Zancheta R, Paro AH, Possi AP, Cruz FC and Planeta CS, 2011. Comparison of
716 caffeine-induced locomotor activity between adolescent and adult rats. Eur J Pharmacol,
717 660:363-367. doi: 10.1016/j.ejphar.2011.03.052

718 Moser V, Bailey F, Bowers W, Raffaele K, Crofton K and Gilbert M, 2016. NAFTA (North
719 American Free Trade Agreement) Technical Working Group on Pesticides (TWG),
720 Developmental Neurotoxicity Study (DNT) Guidance Document

721 Montévil M, Acevedo N, Schaeberle CM, Bharadwaj M, Fenton SE and Soto AM, 2020. A
722 Combined Morphometric and Statistical Approach to Assess Nonmonotonicity in the
723 Developing Mammary Gland of Rats in the CLARITY-BPA Study. Environmental Health
724 Perspectives, 128:057001. doi: doi:10.1289/EHP6301

725 National Academies of Sciences, Engineering, and Medicine, 2017. Application of Systematic
726 Review Methods in an Overall Strategy for Evaluating Low-Dose Toxicity from Endocrine
727 Active Chemicals. Washington, DC: The National Academies Press.
728 https://doi.org/10.17226/24758.

729 Perharič L, Juvan KA and Stanovnik L, 2013. Acute effects of a low-dose atropine/scopolamine
730 mixture as a food contaminant in human volunteers. Journal of Applied Toxicology, 33:980-
731 990

732 Picciotto MR, Addy NA, Mineur YS and Brunzell DH, 2008. It is not "either/or": activation and
733 desensitization of nicotinic acetylcholine receptors both contribute to behaviors related to
734 nicotine addiction and mood. Prog Neurobiol, 84:329-342. doi:
735 10.1016/j.pneurobio.2007.12.005

Pitschner HF and Wellstein A, 1988. Dose-response curves of pirenzepine in man in relation to M1- and M2-cholinoceptor occupancy. Naunyn-Schmiedeberg's Archives of Pharmacology, 338:207-210. doi: 10.1007/BF00174872

Pitschner HF, Schulte B, Neuzner J, Wellstein A, Palm D and Schlepper M, 1994. [Subtypes of muscarinic receptors--aspects of their physiologic significance for controlling heart rate in the human]. Z Kardiol, 83 Suppl 5:9-20

Puatanachokchai R, Morimura K, Wanibuchi H, Oka M, Kinoshita A, Mitsuru F, Yamaguchi S, Funae Y and Fukushima S, 2006. Alpha-benzene hexachloride exerts hormesis in preneoplastic lesion formation of rat hepatocarcinogenesis with the possible role for hepatic detoxifying enzymes. Cancer Letters, 240:102-113. doi: https://doi.org/10.1016/j.canlet.2005.09.006

Shutoh Y, Takeda M, Ohtsuka R, Haishima A, Yamaguchi S, Fujie H, Komatsu Y, Maita K and Harada T, 2009. Low dose effects of dichlorodiphenyltrichloroethane (DDT) on gene transcription and DNA methylation in the hypothalamus of young male rats: implication of hormesis-like effects. J Toxicol Sci, 34:469-482. doi: 10.2131/jts.34.469

Sukata T, Uwagawa S, Ozaki K, Ogawa M, Nishikawa T, Iwai S, Kakehashi A, Wanibuchi H, Imaoka S, Funae Y, Okuno Y and Fukushima S, 2002. Detailed low-dose study of 1,1-BIS(p-chlorophenyl)-2,2,2- trichloroethane carcinogenesis suggests the possibility of a hormetic effect. International journal of cancer. Journal international du cancer, 99:112-118. doi: 10.1002/ijc.10312

Takeda H, Tsuji M, Miyamoto J and Matsumiya T, 2002. Rosmarinic acid and caffeic acid reduce the defensive freezing behavior of mice exposed to conditioned fear stress. Psychopharmacology, 164:233-235. doi: 10.1007/s00213-002-1253-5

Taylor JA, Shioda K, Mitsunaga S, Yawata S, Angle BM, Nagel SC, Vom Saal FS and Shioda T, 2018. Prenatal Exposure to Bisphenol A Disrupts Naturally Occurring Bimodal DNA Methylation at Proximal Promoter of fggy, an Obesity-Relevant Gene Encoding a Carbohydrate Kinase, in Gonadal White Adipose Tissues of CD-1 Mice. Endocrinology, 159:779-794. doi: 10.1210/en.2017-00711

Tyl RW, Myers CB, Marr MC, Thomas BF, Keimowitz AR, Brine DR, Veselica MM, Fail PA, Chang TY, Seely JC, Joiner RL, Butala JH, Dimond SS, Cagen SZ, Shiotsuka RN, Stropp GD and Waechter JM, 2002. Three-generation reproductive toxicity study of dietary bisphenol A in CD Sprague-Dawley rats. Toxicol Sci, 68:121-146. doi: 10.1093/toxsci/68.1.121

Uchtmann KS, Taylor JA, Timms BG, Stahlhut RW, Ricke EA, Ellersieck MR, vom Saal FS and Ricke WA, 2020. Fetal bisphenol A and ethinylestradiol exposure alters male rat urogenital tract morphology at birth: Confirmation of prior low-dose findings in CLARITY-BPA. Reproductive Toxicology, 91:131-141. doi: https://doi.org/10.1016/j.reprotox.2019.11.007

Varlinskaya EI and Spear LP, 2009. Ethanol-Induced Social Facilitation in Adolescent Rats: Role of Endogenous Activity at Mu Opioid Receptors. Alcoholism: Clinical and Experimental Research, 33:991-1000. doi: 10.1111/j.1530-0277.2009.00920.x

Wellstein A and Pitschner HF, 1988. Complex dose-response curves of atropine in man explained by different functions of M1- and M2-cholinoceptors. Naunyn-Schmiedeberg's Archives of Pharmacology, 338:19-27. doi: 10.1007/BF00168807

Wildemann TM, Mirhosseini N, Siciliano SD and Weber LP, 2015. Cardiovascular responses to lead are biphasic, while methylmercury, but not inorganic mercury, monotonically increases blood pressure in rats. Toxicology, 328:1-11. doi: 10.1016/j.tox.2014.11.009

Yuanqing H, Suhua W, Guangwei X, Chunlan R, Hai Q, Wenrong X, Rongzhu L, Aschner M and Milatovic D, 2013. Acrylonitrile has Distinct Hormetic Effects on Acetylcholinesterase Activity in Mouse Brain and Blood that are Modulated by Ethanol. Dose-Response, 11:dose-response.11-030.Yuanqing. doi: 10.2203/dose-response.11-030.Yuanqing

Zhang Q, Yu Y-P, Ye Y-L, Zhang J-T, Zhang W-P and Wei E-Q, 2011. Spatiotemporal properties of locomotor activity after administration of central nervous stimulants and sedatives in mice. Pharmacology Biochemistry and Behavior, 97:577-585. doi: https://doi.org/10.1016/j.pbb.2010.09.011

Zhang Y, Lu R, Liu W, Wu Y, Qian H, Zhao X, Wang S, Xing G, Yu F and Aschner M, 2012. Hormetic Effects of Acute Methylmercury Exposure on GRP78 Expression in Rat Brain Cortex. Dose-Response, 11:dose-response.11-055.Rongzhu. doi: 10.2203/dose-response.11-055.Rongzhu

Zorrilla LM, Gibson EK, Jeffay SC, Crofton KM, Setzer WR, Cooper RL and Stoker TE, 2009. The Effects of Triclosan on Puberty and Thyroid Hormones in Male Wistar Rats. Toxicological Sciences, 107:56-64. doi: 10.1093/toxsci/kfn225

Zsarnovszky A, Le HH, Wang HS and Belcher SM, 2005. Ontogeny of rapid estrogen-mediated extracellular signal-regulated kinase signaling in the rat cerebellar cortex: potent nongenomic agonist and endocrine disrupting activity of the xenoestrogen bisphenol A. Endocrinology, 146:5388-5396. doi: 10.1210/en.2005-0565

Abbreviations

AOP	adverse outcome pathway
AROI	acceptable range of oral intake
BPA	Bisphenol A
CEP	EFSA Panel on Food Contact Materials, Enzymes and Processing Aids
CLARITY-BPA	Consortium Linking Academic and Regulatory Insights on BPA Toxicity
COX	cyclooxygenase
CP	checkpoint
DBP	Dibutyl phthalate
DEHP	Bis(2-ethylhexyl) phthalate
DNT	developmental neurotoxicity
ECHA	European Chemicals Agency
EMA	European Medicines Agency
EUROTOX	Federation of European Toxicologists and European Societies of Toxicology
F2	second filial generation
FAO	Food and Agriculture Organization of the United Nations
FDA	Food and Drug Administration
HBGV	health-based guidance values
HPG	hypothalamic-pituitary-gonadal axis
ip	intraperitoneal
IUTOX	International Union of Toxicology
JRC	Joint Research Centre
MDR	monotonic dose-response
MOA	mode of action
MOE	margin of exposure
NAMs	new approach methodologies
nAChRs	nicotinic acetylcholine receptors
NMDR	non-monotonic dose-response
NMDRC	non-monotonic dose-response curve
NOAEL	no observed adverse effect level
OECD	Organisation for Economic Co-operation and Development
P_{NMDR}	probability of non-monotonic dose-response
PCBs	Polychlorinated Biphenyls
PGE2	Prostaglandin E2
PND	post-natal day
RA	risk assessment
RP	reference point
SC	Scientific Committee

sc	subcutaneous
SR	systematic review
T	testosterone
ToR	Terms of Reference
US EPA	United States Environmental Protection Agency
WHO	World Health Organization

820

821

Appendix A. Results from the additional probabilistic assessments

The Table A and B show the results of additional probabilistic assessments for Perharic et al., 2013 conducted according to the methodology proposed by Chevillotte et al. (2017a,b).

Table A: From Perharic et al., 2013 (Table 4). Endpoint: Heart rate at 2h

	Probabilistic methodology Monte-carlo resampling	**Probabilistic methodology Latin-Hypercube resampling**
Type of Dose-Response	Prob,(%)	Prob,(%)
No DR	0	0
MDR increasing	33.8	0.02
MDR decreasing	0	0
NMDR U	66.1	99.98
NMDR inverted-U	0	0
NMDR complex	0.06	0
Total	100	100

Table B: From Perharic et al., 2013 (Table 4). Endpoint: Heart rate at 2.75h

	Probabilistic methodology Monte-carlo resampling	**Probabilistic methodology Latin-Hypercube resampling**
Type of Dose-Response	Prob,(%)	Prob,(%)
No DR	0	0
MDR increasing	11.34	0
MDR decreasing	0	0
NMDR U	86.7	100
NMDR inverted-U	0	0
NMDR complex	1.96	0
Total	100	100

The Table C, D and E show the results of additional probabilistic assessments for Uchtmann et al., 2020 conducted according to the methodology proposed by Chevillotte et al. (2017a,b).

Table C: From Uchtmann et al., 2020 (Table 3 - Supplementary material). Endpoint: Body weight (litter) at PND1

	Probabilistic methodology Monte-carlo resampling	**Probabilistic methodology Latin-Hypercube resampling**
Type of Dose-Response	Prob,(%)	Prob,(%)
No DR	3.7	0.39
MDR increasing	25.1	53.7
MDR decreasing	4.95	0.01
NMDR U	0.9	0
NMDR inverted-U	58.8	45.9
NMDR complex	6.5	0
Total	100	100

Table D: From Uchtmann et al., 2020 (Table 3 - Supplementary material). Endpoint: Colliculus angle (litter) at PND1

	Probabilistic methodology Monte-carlo resampling	Probabilistic methodology Latin-Hypercube resampling
Type of Dose-Response	Prob,(%)	Prob,(%)
No DR	8.9	6.25
MDR increasing	3.1	0
MDR decreasing	47.2	93.7
NMDR U	33.73	0.07
NMDR inverted-U	1	0
NMDR complex	6.13	0
Total	100	100

835

836 **Table E**: From Uchtmann et al., 2020 (Table 3 - Supplementary material). Endpoint: urogenital
837 sinus epithelium thickness (midway section)

	Probabilistic methodology Monte-carlo resampling	Probabilistic methodology Latin-Hypercube resampling
Type of Dose-Response	Prob,(%)	Prob,(%)
No DR	14.73	14.43
MDR increasing	3.23	0
MDR decreasing	49.92	85.57
NMDR U	31.34	0
NMDR inverted-U	0.11	0
NMDR complex	0.67	0
Total	100	100

838

839

840

841

842

843

844

845

846

847

848

849

850

851

Annex I. Assessment of non-monotonicity claims for BPA

In the external report on NMDR (Beausoleil et al., 2016), BPA is reported as the substance under the EFSA remit with the highest number of *in vivo* datasets for which the authors report a potential NMDR (35). BPA was also identified in the targeted literature search conducted for this assessment for updating the information. Once characteristic of these studies are indications of NMDR present at relatively low dose BPA exposure, which have been claimed for several non-apical endpoints (Lagarde et al, 2015). One limitation of many of these studies is use of two or three dose groups (in addition to controls), which is no well-suited to assess the presence of NMDR with any reasonable certainty.

Claims of NMDR have also made in several publications based on data from the Consortium Linking Academic and Regulatory Insights on BPA Toxicity (CLARITY-BPA) Program conducted by the US National Toxicology Program. The studies linked to the CLARITY-BPA Program cover the wide range of BPA doses with reported effects in the scientific literature, and a large number of effects measured by different groups from animal samples with identical BPA exposure conditions, generated by the same facility in FDA's National Center for Toxicological Research (NCTR). The participant laboratories received blinded samples, meaning they did not know whether samples had been dosed with BPA or how much, to minimize the potential for bias. Consequently, these studies very considered particularly relevant for addressing NMDR claims for BPA, and were added to those retrieved in the literature search.

This annex covers exclusively the evaluation of the reliability of the NMDR claims for several publication identified by the Report and the targeted search done for this assessment. One aim of this exercise is to support the EFSA risk assessment on BPA by the CEP Panel.

Table AI-1. Studies on BPA with datasets on NMDR included in the Report by Beausoleil et al. (2016), and additional studies including those from the BPA-Clarity program assessed for NMDR.

Publication, chemical, and measured effects	**Dose range, # of dose-groups (N) excluding controls**	**1. Presence/shape of NMDR** (checkpoints not fulfilled*)	**2. Nature of measured effect**	**3. Biol plaus***	**4. role in adversity***	**5. Probability of NMDR (%) as described by Chevillotte et al. (2017a,b)**	**Comments**
Studies identified in the Report							
Tyl et al., 2002, **BPA**, three generation reproductive toxicity study in rats. 1. Absolute liver weight in F2 females	0.001-500 mg/kg bw day in the diet N=6	1. No (CP-3 and CP-5)	1. Intermediate	1. Yes	1.Yes, increase in liver weight may be indicative of possible adverse effects; however in this study no histopathological changes in liver were observed for this group.	1. P_{NMDR} 66 (U)	If there is a NMDR then it is driven by one dose group (no clear trend in the surrounding dose groups that may explain NMDR). NMDR was assessed for other effects: relative liver weight, paired testes weight, and anogenital distance in F2 females, but met only 3 or less checkpoints
Zsarnovszky et al., 2005, *in vivo* and *in vitro* effects of **BPA**, 17β-estradiol (E2) and their mixture on cereberal signally in rats. 1.Extracellular kinase signaling in cerebelar cortex: pERK-IRCellAtP10	Intracerebelar injection of 3 µL per animal of BPA concentrations 10^{-12} to 10^{-6} M N=7	1. Yes ∩, second increase observed at the highest doses (only 3 checkpoints met)	1. Intermediate	1. Yes	1. Unclear	1. P_{NMDR} 100 (complex)	E2 at the same doses and conditions provokes the same NMDR response, even in quantitative terms, suggesting equipotency for E2 and BPA. Co-injection of E2 and BPA inhibits the response A parallel *in vitro* study on primary cerebellar granule cells, range 10^{-12} to 10^{-4} M, N=5, reported ∩ shape response for induction of ERK phosphorylation
Angle et al., 2013, effects of in utero **BPA** exposure in mice 1. Gonadal fat pad weight	0.005-50 mg/kg bw day in the diet N=5	1. Yes ∩ (CP-3 and CP-5) 2. Yes ∩ (CP-3 and CP-5) 3. Yes ∩	1.Intermediate 2.Intermediate 3.Intermediate	1.Yes 2. Yes 3.Yes	1, 2 and 3 Yes but what effect size?	1. P_{NMDR} 79 (complex) 2. P_{NMDR} 99 (U)	Some departure form monotonicity seems present but random fluctuation in response also plausible

2 Renal fat pad weight 3 Serum adiponectin		(CP-2 and CP-3)				3. P_{NMDR} 35 (U)	The Report also include dataset for other endpoints, fulfilling 3 or less checkpoints
Kendig et al., 2012, estrogen-like effects of in utero **BPA** or 17α-ethilyl estradiol (EE) exposure in mice 1. Sperm count 2. Sperm motility	0.004-40 mg/kg bw day in the diet N=5	1, Not (CP-3 and CP-6) 2. Yes ∩ (CP-3 and CP-6)	2. Intermediate	2. Yes	2. Unclear, can be considered beneficial?	1. P_{NMDR} 35 (U) 2. P_{NMDR} 58.44	Findings are inconsistent with (Hass et al., 2016) and findings from the Clarity study (Clarity BPA, NTP 2018) Similar shape may be seen for EE but difficult to assess as it is based on only 3 doses
Studies not included in the Report							
Hass et al., 2016, effect of **BPA** in utero exposure in rats 1.Sperm count in male offspring (Figure 2) 2. Swim length of female offspring (Figure 4A)	0.025-50mg/kg bw day by gavage N=4	1. Yes U 2. Yes U	1. Intermediate 2. Apical	1. Yes 2. Yes	1. Yes 2. Yes		1. Modest effect (less than 20% reduction *vs.* control). Similar NMDR not observed in a comparable study (Kendig et al., 2012) or the Clarity study (Clarity BPA, NTP 2018). 2. Again modest effect (less than 20% reduction vs. control) for swim length. Also a U shape for males, but at different dose levels and differences are not statistically significant
Taylor et al., 2018, effects of **BPA** prenatal exposure in mice 1. Gonadal fat pads weight (Figure 1 B)	0.005-0.5 mg/kg bw per day by gavage N=2	1 Yes ∩ but only control and 2 dose groups. Flattens out for males	1. Intermediate	1.Yes	1. Yes but what effect size?		A control and 2 doses are not suitable for evaluating NMDR but dose range and pattern is in line with findings reported in Angle et al., 2013 above.
Dere et al., 2018 (Clarity) effects of **BPA** early gestation exposure in rats,	0.0025-250 mg/kg bw day by gavage N=6	1.Yes ∩	1. Early effect	1. Yes	1. Unclear as no effects are observed on semen quality in the clarity study		

1 Sperm DNA methylation (Figure 2)					(Clarity BPA, NTP 2018)		
Badding et al., 2019 (Clarity), effects of **BPA** early gestation exposure in rats. This paper evaluated NMDR using the six checkpoints for all outcomes with suspected NMDR. Authors identified: 1. Percent basophils at 1-y in stop arm for females (Table 4, Figures 1 and 2) 2. Total bile acids at 1-y in stop arm for males (Table 5) Authors discarded other outcomes as unlikely (<5 checkpoints)	0.0025-25 mg/kg bw day by gavage N=5	1. Yes ∩ 2. Yes U	1. Intermediate 2. Intermediate	1. Yes? 2. Yes?	1 and 2 unclear?		NMDR seems quite clear but replication in another study would strengthen these findings. Biological relevance is unclear (to be specifically checked with the BPA group)
Uchtmann et al., 2020 (Clarity), effects of **BPA** early gestation exposure in rats. 1.Body weight (Figure 4), 2. Fetal urogenital sinus epithelium thickness (Figure 7)	0.0025-25 mg/kg bw day by gavage N=5	1. Unclear 2.Unclear	1. Apical 2. Intermediate?	1. Unclear 2. Unclear	1. For body weight it is unclear what effect size in rodents is biologically relevant 2. Same for urogenital sinus	1. P_{NMDR} 58.8 2. P_{NMDR} 31 (U)	1. High variability within dose groups. Lack of NMDR at all other postnatal dates 2. Absence of adverse effect on female reproductive outcomes leaves a question mark on the biological relevance of the findings on urogenital sinus.

<u>Li et al., 2019</u> (literature search), **e**ffects of BPA, peri/post pubertal exposure in male mice 1.SREBP-1c mRNA/protein expression (Figure 3) 2.SREBP-2 mRNA/protein expression (Figure 1) 3.HMGCR mRNA/protein expression (Figure 1) 4.SCD-1 mRNA/protein expression (Figure 3) 5.Serum triglycerides and total cholesterol (Table 4) 6.Serum LDL-C, HDL-C, ALT, AST (Table 4) 7.Liver triglycerides and total cholesterol (Table 4)	0.05- 5 mg/kg bw day in the diet N=3	1. Yes ∩ 2. Yes ∩ 3. Yes ∩ 4. Yes ∩ 5. Yes ∩ 6. Yes ∩ 7. Yes ∩	All early effects	All unclear	All mechanistic information not relevant, in isolation, for the consideration of adverse	6. P_{NMDR} 98.4 for ALT 7. P_{NMDR} 70.1	Significant differences at 0.05 and 0.5 mg/kg bw day but not at 5 mg/kg bw day. Changes in biochemical parameters are very small, and it is not mentioned whether they are within the historical control range.

Rubin et al., 2017 (literature search), effects of BPA, perinatal or perinatal and peripubertal exposure in mice 1.Body weight in female exposed perinatally and peripubertally (Figure 1) 2. Body composition in female exposed perinatally and peripubertally (Figure 6)	0.00025-0.250 mg/kg bw/day subcutaneous exposure perinatally and by drinking water peripubertally N=4	1. No 2. No	1. Apical 2. All Intermediate				Only two doses and control, not suitable for NMDR assessment
Yang et al., 2016. (literature search), **e**ffects of BPA, pubertal exposure in mice 1.Body weight (Figure 1) 2.Fat mass (Figure 1) 3.iWAT and eWAT (Figure 1) 4.C/EBP- α (Figure 3) 5.SREBP-1c (Figure 3) 6.SCD-1 (Figure 3) 7.Inflammation (Figure 5) Effects of BPA metabolites on humans 1.Plasma Leptin in lean female subjects (Figure 6) 2.TNFα levels in lean female subjects (Figure 6)	0.0005-5 mg/kg bw day in the diet N=4	No for body weight and fat mass. Unclear for other measures	Early to intermediate (?)		Unclear		Changes in body weight and fat mass are randomly distributed. All other effects are very early events providing mechanistic information and are not used as RP in risk assessment. They are seen only at highest dose, maybe due to overt toxicity (100xthe TDI)

Sharma et al., 2019 (literature search), effects of BPA, exposure in mice 1.PPAR (α, β, γ) mRNA 2. protein expression in testes (Figure 2)	4-16 mg/kg/day intraperitoneally N=3	1. No 2 Yes ∩	1. Early event 2. Early event		Unclear		1. Monotonic decrease in all dose groups, however controls were lower than the lowest dose group. The apical effect (pattern of histopathological effects) was monotonic.
Zhang et al., 2019 (literature search), human cohort study of pregnant women 1.Fasting plasma glucose (Figure 1)	Urine samples collected at ~13 weeks of gestation to examine the concentration of 4 bisphenols (BPA, BPS, BPF, BPAF)	1. Yes U	1. Intermediate		Unclear		NMDR (U-shaped curve) observed only in fasting plasma glucose levels among overweight pregnant women. For overweight women higher BPA concentrations were, however, associated with lower risk of GDM. This association is inconsistent with the pattern observed for fasting plasma glucose levels (based on the NDMR for fasting plasma glucose one would expect to see higher risk of GDM at high BPA exposures). As such these findings appear inconsistent
Zhou et al., 2017 (literature search), effects on BPA, pubertal exposure in male mice (n=8, 8week exposure) 1.Neuron quantity in the CA3 region of the hippocampus (Figure 4)	0.0005-5 mg/kg bw /day by gavage N= 3	1. No	1. Intermediate				Decrease in low- and high-dose group. No effect in mid-dose group. In another region of the hippocampus there was no effect on the neuron quantity and in a third region there was a decrease in the high-dose group.

*CP = checkpoint as defined in the Report (see full list at the "Introduction" section):
CP-3. Can the apparent NMDR be explained by one single potential outlying dose group?
CP-5. Is the steepness of the dose-response curve outside the range of biologically plausible/realistic dose-response shapes?
∏ The symbol U indicates a NMDR with U (or J) shape, the symbol ∩ indicates a NMDR with inverted U (or J) shape
‡ Only addressed when a possible NMDR is confirmed under 1. Presence/ shape of NMDR

In the EFSA external report on NMDR (EFSA 2015) four studies on BPA were identified and evaluated with respect to the six checkpoints (Tyl et al., 2002, Zsarnovszky et al., 2005, Kendig 2012 et al., and Angle et al., 2013). A U-shaped NMDR was identified for liver weight in the F2 generation, intracellular signalling (pERK-IRCellAtP10) and cell numbers in gonadal and renal fat pads; while an inverse U-shaped NMDR was observed for semen quality. Each of these studies only fulfilled 4 checkpoints or less. Independent review of these studies in Table AI-1 are in line with the Report that the presence of NMDR is subject to some uncertainty. The six check points are, however, primarily based on statistical considerations for evaluating a single study and they do not address accumulated evidence from more than one study. To address this uncertainty, outcomes included in the report were addressed further by screening for more recent studies that may conform these findings. No studies on liver weight or intracellular signalling (pERK-IRCellAtP10) were identified. For sperm count, Hass et al. (2016) reported a U-shaped association with sperm quality, which is in opposite direction with the NMDR reported by Kendig et al. (2012). In the more recent Clarity study (Clarity BPA, NTP 2018), no indications of NMDR were observed. Overall findings on NMDR and male fertility appear inconsistent.

Using data form the CLARITY study Montevil et al. (2020) identified, using advanced statistical methods, an NMDR between developmental exposure and offspring mammary gland development. The observed NMDR was rather unconventional with the slope changing sign two times. Such a pattern (increase followed by decrease and again increase or the reverse) is quite unique and difficult to compare with other studies in the context of regulatory risk assessment. In the absence of any clear biological explanation why the dose response curve may behave in such non-linear manner and taking into consideration lack of overall significance (from the NULL model) the pattern observed may be a result of overfitting of the data rather than a true biological relationship. In any case the findings from this paper need to be replicated before any conclusions on relevance and adversity can be made.

Using the six checkpoints Badding et al. (2019) identified NMDR for %basophils for females and total bile acids at 1-y in stop arm for males in the Clarity study. Similar findings have not been reported in previous studies. Finally, the presence of NMDR following in utero exposure has been observed in some but not all studies on BPA (Lagarde et al., 2015). These findings may be in line with findings on NMDR for cell numbers in renal and gonadal fat pads (Angle et al., 2013). Overall findings on NMDR for weight appear unstable and they may be sensitive to various experimental conditions (Lagarde et al., 2015). The relevance of such possible NMDR is perhaps best highlighted in the Clarity study where some indications of NMDR at postnatal day 1 has been claimed (Uchtmann et al., 2020). Even if so no further difference in weight between dose groups was observed at later time points (Clarity BPA, NTP 2018) making the biological relevance of this observation highly uncertain. In summary the endpoints identified and consistency of findings across studies do not suggest that NMDR is of relevance for the risk assessment of BPA.

Answer to the questions (proposed approach)

What is the experimental evidence for the effect observed (*in vitro* /*in vivo*? Other?)

- There is a number of *in vivo* studies claiming NMDR for some early, intermediate and apical effects. The lack of consistency among results triggered the Clarity study.
- Statistical assessments have identified some NMDR datasets extracted from the Clarity study, e.g. for weight at specific time points. However, for each outcome there is a lack of consistency across existing studies.

What is the biological relevance of the effects observed? Can a (quantitative) relation between the observed effect and an adverse outcome be established? Ideally: Could a mechanistic sequence (AOP) be partially or fully established? If yes, specific considerations need to be applied and a diversion from the current methodologies for RA may be needed

- In addition to the lack of consistency in the findings claiming NMDRs across studies, for several outcomes where NMDR has been claimed no biological explanation connecting mechanistically the claimed NMDR has been established.
- Monotonic responses are observed for those endpoints relevant for establishing the RP
- The assessment does not suggest that NMDR is of relevance for the risk assessment of BPA.

If information is lacking on whether an observed effect can lead to an adverse outcome, additional testing may be needed. Here NAMs would be of relevance given the need for identifying a mechanistic sequence of events.

References

Angle BM, Do RP, Ponzi D, Stahlhut RW, Drury BE, Nagel SC, Welshons WV, Besch-Williford CL, Palanza P, Parmigiani S, vom Saal FS and Taylor JA, 2013. Metabolic disruption in male mice due to fetal exposure to low but not high doses of bisphenol A (BPA): evidence for effects on body weight, food intake, adipocytes, leptin, adiponectin, insulin and glucose regulation. Reprod Toxicol, 42:256-268. doi: 10.1016/j.reprotox.2013.07.017

Badding MA, Barraj L, Williams AL, Scrafford C and Reiss R, 2019. CLARITY-BPA Core Study: Analysis for non-monotonic dose-responses and biological relevance. Food Chem Toxicol, 131:110554. doi: 10.1016/j.fct.2019.06.001

Beausoleil C, Beronius A, Bodin L, Bokkers BGH, Boon PE, Burger M, Cao Y, De Wit L, Fischer A, Hanberg A, Leander K, Litens-Karlsson S, Rousselle C, Slob W, Varret C, Wolterink G and Zilliacus J, 2016. Review of non-monotonic dose-responses of substances for human risk assessment. EFSA Supporting Publications, 13:1027E. doi: 10.2903/sp.efsa.2016.EN-1027

Dere E, Anderson LM, Huse SM, Spade DJ, McDonnell-Clark E, Madnick SJ, Hall SJ, Camacho L, Lewis SM, Vanlandingham MM and Boekelheide K, 2018. Effects of continuous bisphenol A exposure from early gestation on 90 day old rat testes function and sperm molecular profiles: A CLARITY-BPA consortium study. Toxicol Appl Pharmacol, 347:1-9. doi: 10.1016/j.taap.2018.03.021

Hass U, Christiansen S, Boberg J, Rasmussen MG, Mandrup K and Axelstad M, 2016. Low-dose effect of developmental bisphenol A exposure on sperm count and behaviour in rats. Andrology, 4:594-607. doi: 10.1111/andr.12176

Kendig EL, Buesing DR, Christie SM, Cookman CJ, Gear RB, Hugo ER, Kasper SN, Kendziorski JA, Ungi KR, Williams K and Belcher SM, 2012. Estrogen-like disruptive effects of dietary exposure to bisphenol A or 17α-ethinyl estradiol in CD1 mice. Int J Toxicol, 31:537-550. doi: 10.1177/1091581812463254

Lagarde F, Beausoleil C, Belcher SM, Belzunces LP, Emond C, Guerbet M and Rousselle C, 2015. Non-monotonic dose-response relationships and endocrine disruptors: a qualitative method of assessment. Environ Health, 14:13. doi: 10.1186/1476-069x-14-13

Li Q, Zhang H, Zou J, Mai H, Su D, Feng X and Feng D, 2019. Bisphenol A exposure induces cholesterol synthesis and hepatic steatosis in C57BL/6 mice by down-regulating the DNA methylation levels of SREBP-2. Food Chem Toxicol, 133:110786. doi: 10.1016/j.fct.2019.110786

Montévil M, Acevedo N, Schaeberle CM, Bharadwaj M, Fenton SE and Soto AM, 2020. A Combined Morphometric and Statistical Approach to Assess Nonmonotonicity in the Developing Mammary Gland of Rats in the CLARITY-BPA Study. Environmental Health Perspectives, 128:057001. doi: doi:10.1289/EHP6301

NTP (National Toxicology Program), 2018. NTP Research Report on the CLARITY-BPA Core Study: A Perinatal and Chronic Extended-Dose-Range Study of Bisphenol A in Rats.

Rubin BS, Paranjpe M, DaFonte T, Schaeberle C, Soto AM, Obin M and Greenberg AS, 2017. Perinatal BPA exposure alters body weight and composition in a dose specific and sex

specific manner: The addition of peripubertal exposure exacerbates adverse effects in female mice. Reprod Toxicol, 68:130-144. doi: 10.1016/j.reprotox.2016.07.020

Sharma S, Ahmad S, Afjal MA, Habib H, Parvez S and Raisuddin S, 2019. Dichotomy of bisphenol A-induced expression of peroxisome proliferator-activated receptors in hepatic and testicular tissues in mice. Chemosphere, 236:124264. doi: 10.1016/j.chemosphere.2019.06.234

Taylor JA, Shioda K, Mitsunaga S, Yawata S, Angle BM, Nagel SC, Vom Saal FS and Shioda T, 2018. Prenatal Exposure to Bisphenol A Disrupts Naturally Occurring Bimodal DNA Methylation at Proximal Promoter of fggy, an Obesity-Relevant Gene Encoding a Carbohydrate Kinase, in Gonadal White Adipose Tissues of CD-1 Mice. Endocrinology, 159:779-794. doi: 10.1210/en.2017-00711

Tyl RW, Myers CB, Marr MC, Thomas BF, Keimowitz AR, Brine DR, Veselica MM, Fail PA, Chang TY, Seely JC, Joiner RL, Butala JH, Dimond SS, Cagen SZ, Shiotsuka RN, Stropp GD and Waechter JM, 2002. Three-generation reproductive toxicity study of dietary bisphenol A in CD Sprague-Dawley rats. Toxicol Sci, 68:121-146. doi: 10.1093/toxsci/68.1.121

Uchtmann KS, Taylor JA, Timms BG, Stahlhut RW, Ricke EA, Ellersieck MR, vom Saal FS and Ricke WA, 2020. Fetal bisphenol A and ethinylestradiol exposure alters male rat urogenital tract morphology at birth: Confirmation of prior low-dose findings in CLARITY-BPA. Reproductive Toxicology, 91:131-141. doi: https://doi.org/10.1016/j.reprotox.2019.11.007

Wellstein A and Pitschner HF, 1988. Complex dose-response curves of atropine in man explained by different functions of M1- and M2-cholinoceptors. Naunyn-Schmiedeberg's Archives of Pharmacology, 338:19-27. doi: 10.1007/BF00168807

Yang M, Chen M, Wang J, Xu M, Sun J, Ding L, Lv X, Ma Q, Bi Y, Liu R, Hong J and Ning G, 2016. Bisphenol A Promotes Adiposity and Inflammation in a Nonmonotonic Dose-response Way in 5-week-old Male and Female C57BL/6J Mice Fed a Low-calorie Diet. Endocrinology, 157:2333-2345. doi: 10.1210/en.2015-1926

Zhang W, Xia W, Liu W, Li X, Hu J, Zhang B, Xu S, Zhou Y, Li J, Cai Z and Li Y, 2019. Exposure to Bisphenol a Substitutes and Gestational Diabetes Mellitus: A Prospective Cohort Study in China. Front Endocrinol (Lausanne), 10:262. doi: 10.3389/fendo.2019.00262

Zhou YX, Wang ZY, Xia MH, Zhuang SY, Gong XB, Pan JW, Li CH, Fan RF, Pang QH and Lu SY, 2017. Neurotoxicity of low bisphenol A (BPA) exposure for young male mice: Implications for children exposed to environmental levels of BPA. Environmental Pollution, 229:40-48. doi: 10.1016/j.envpol.2017.05.043

Zsarnovszky A, Le HH, Wang HS and Belcher SM, 2005. Ontogeny of rapid estrogen-mediated extracellular signal-regulated kinase signaling in the rat cerebellar cortex: potent nongenomic agonist and endocrine disrupting activity of the xenoestrogen bisphenol A. Endocrinology, 146:5388-5396. doi: 10.1210/en.2005-0565

Abbreviations

AOP	adverse outcome pathway
BPA	Bisphenol A
CEP	EFSA Panel on Food Contact Materials, Enzymes and Processing Aids
CLARITY-BPA	Consortium Linking Academic and Regulatory Insights on BPA Toxicity
CP	checkpoint
E2	estradiol
F2	second generation
FDA	Food and Drug Administration
NAMs	new approach methodologies
NMDR	non-monotonic dose-response
P_{NMDR}	probability of non-monotonic dose-response
RP	reference point
TDI	tolerable daily intake

Annex II. Assessment of non-monotonicity claims for phthalates

Introduction

In the external report on NMDR (Beausoleil et al. 2016) phthalates (DEHP and DBP) are reported with the substances under EFSA remit with the highest number of *in vivo* datasets reporting potential NMDR (30 for DEHP and 5 for DBP). For one data set, aromatase activity in rats exposed to DEHP (Andrade et al., 2006), the six checkpoints were met. Phthalates in general and DEHP in particular were also identified in the targeted literature search conducted for updating the information. Consequently, specific assessments of NMDR have been considered in this opinion. This Annex presents the assessment for the phthalates, focusing on DEHP.

Key elements from the EFSA assessment on phthalates

Phthalates are plasticizers used as FCM under the EFSA domain. Several phthalates are considered as having ED properties, are classified as toxic for the reproduction (CLP 1B), considered substances of very high concern (SVHC) requiring authorization prior to use (Annex XIV) and have use restrictions (Annex XVII) under the REACH Regulation.

The EFSA CEP Panel established a temporary group-TDI of 50 µg/kg bw/day for four phthalates (dibutyl phthalate (DBP), benzyl butyl phthalate (BBP), bis(2-ethylhexyl) phthalate (DEHP), diisononyl phthalate (DINP). One of the criteria for grouping these phthalates was a common mode of action, reduction in fetal testosterone level as an intermediate key event. In particular, "*with regard to the grouping of these phthalates due to similar reproductive effects, the CEP Panel considered the reduction of the fetal testosterone production during a window of susceptibility in rats induced by DBP, BBP and DEHP as a critical step in the reproductive toxicity of the phthalates. This effect provided the basis for grouping together these phthalates, there being a mechanistic rationale for the plausibility and validity of grouping* (EFSA CEP Panel, 2019[7])".

The reduction of testosterone level is widely recognised as a critical step for the degeneration of androgen-dependent tissues (AOP 288: Collet, 2020) (NAS, 2017[8]). Therefore, the EFSA assessment on phthalates is mainly focused on their reproductive effects, indicating that a full assessment of all other adverse effects was not feasible within the mandate timelines, as elucidated in the Section 1.2. that states "in compliance with the European Commission mandate referring to the predefined dataset underlying the 2017 ECHA's proposal to restrict the use of DBP, BBP, DEHP and DIBP under the REACH Regulation, also this CEP Panel's assessment is mainly centred on phthalate-induced reproductive effects. The CEP Panel is aware of the intrinsic limitations of this approach and considers that all the potential toxicological endpoints should be examined with the same degree of rigour. However, due to

[7] EFSA CEP Panel (EFSA Panel on Food Contact Materials, Enzymes and Processing Aids), Silano V, Barat Baviera JM, Bolognesi C, Chesson A, Cocconcelli PS, Crebelli R, Gott DM, Grob K, Lampi E, Mortensen A, Riviere G, Steffensen I-L, Tlustos C, Van Loveren H, Vernis L, Zorn H, Cravedi J-P, Fortes C, Tavares Pocas MF, Waalkens-Berendsen I, Wolfle D, Arcella D, Cascio C, Castoldi AF, Volk K and Castle L, 2019. Scientific Opinion on the update of the risk assessment of di-butylphthalate (DBP), butyl-benzyl-phthalate (BBP), bis(2 ethylhexyl)phthalate (DEHP), di-isononylphthalate (DINP) and di-isodecylphthalate (DIDP) for use in food contact materials. EFSA Journal 2019;17(12):5838, 85 pp. https://doi.org/10.2903/j.efsa.2019.5838

[8] National Academies of Sciences, Engineering, and Medicine. 2017. Application of Systematic Review Methods in an Overall Strategy for Evaluating Low-Dose Toxicity from Endocrine Active Chemicals. Washington, DC: The National Academies Press. https://doi.org/10.17226/24758.

the limited time for the completion of the opinion and the amount of new evidence available since the 2005 publication of the EFSA Food Additives, Flavourings, Processing Aids and Materials in Contact with Food (AFC) Panel's assessments of DBP, BBP, DEHP, DINP and DIDP (EFSA, 2005a,b,c,d,e), the Panel considered it unfeasible to perform a comprehensive review of all the new data on these phthalates″(EFSA CEP Panel, 2019).

However, the Panel highlighted the concern for other possible effects and "concluded that effects not sufficiently investigated in this opinion, in particular potential effects on neurodevelopment, the immune and/or the metabolic systems for DBP, BBP and DEHP, could be more sensitive endpoints compared to their reproductive toxicity". In particular, regarding neurological and neurodevelopmental effects, the EFSA assessment is in line with the ECHA considerations (2017a)[9] "altered neurodevelopment has been associated with high phthalate exposures in children, as reviewed by Miodovnik et al. (2014). Numerous behavioural disorders including autism spectrum disorders, ADHD, learning disabilities and altered play behaviour have been associated with higher phthalate exposure in humans (reviewed by Braun et al., 2013). Animal studies examining behavioural effects of phthalate exposure have shown some effects that may be related to altered sex differentiation, whereas other behavioural do not appear to be linked with disruption of sex hormones. Different modes of action for phthalate effects on neurodevelopment have been proposed, including interference with the thyroid hormone system, altered calcium signalling, relation to activation of PPARs in brain and altered lipid metabolism (Miodovnik et al., 2014)".

The Panel identified several limitations when evaluating these neurodevelopmental effects and this aspect was considered in the uncertainty analysis and in the recommendations. In particular, in the uncertainty analysis the EFSA CEP Panel mentions that "among several sources of uncertainty identified in a qualitative uncertainty analysis, the main impacts on risk assessment could be attributed to: lack of a sufficient evaluation of toxicity endpoints other than reproduction, i.e. neurodevelopment, immune and/or metabolic system, that could be more sensitive. This could lead to an underestimation of the risk based on the currently proposed group approach focusing on the reproductive effects" (EFSA CEP Panel, 2019).

Data and methodologies

The data source included the studies on phthalates included in the EFSA External Report on NMDR (Beausoleil et al., 2016), complemented with a targeted literature search performed in June 2020 (See Table AII-1 for specifications). In line with the ToRs, the selection focused on *in vivo* mammalian studies and was extended to cover epidemiological studies. The references and citations of the retrieved articles were also searched and relevant studies retrieved and included as results of the search.

The data source was completed with additional information on the effects of phthalates on testosterone levels, obtained from references and citations of the retrieved articles, as well available reports and reviews on phthalates including DEHP and its metabolite MEHP.

The assessment of biological plausibility was based on expert judgement, supported by general knowledge and the specific references mentioned in the assessment section.

[9] ECHA, 2017a. Committee for Risk Assessment (RAC) and Committee for Socio-economic Analysis (SEAC). Opinion on an Annex XV dossier proposing restrictions on four phthalates (DEHP, BBP, DBP, DIBP). ECHA, 2017b. Committee for Risk Assessment (RAC) and Committee for Socio-economic Analysis (SEAC). Background document to the Opinion on the Annex XV dossier proposing restrictions on four phthalates (DEHP, BBP, DBP, DIBP)

Table AII-1 Characteristics and results of the targeted literature search

Database	String	Complementary search	Results
Web of Science selecting the following indexes: SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC.	TS=(nonmonotonic OR monotonic OR non-monotonic OR hormesis OR hormetic OR biphasic OR (nonlinear OR non-linear) OR (inverted AND (curve* OR shape*))) AND TS= (phthalate* OR dehp OR mehp)	The search was complemented with the analysis of the references and citations of the retrieved publications	332 articles retrieved 31 studies selected as final result after the screening

Assessment

The Report (Beausoleil et al. 2016) included two publications on DBP and six on DEHP, the evaluation of the DEHP publications indicated that those from Andrade et al. and Grande et al. corresponded to the same study, and identified one additional publication from the same study not included in the Report, that was added for completeness. The NMDR claims observed in these publications are summarised in Table AII-2.

Table AII-2. Studies on phthalates with datasets on NMDR included in the Report by Beausoleil et al. (2016)

Publication, chemical, and measured effects	**Dose range, # of dose-groups (N) excluding controls**	**1. Presence/shape of NMDR** (checkpoints not fulfilled*)	**2. Nature of measured effect**	**3. Biol plaus***	**4. role in adversity***	**5. Probability of NMDR (%) as described by Chevillotte et al. (2017a,b)**	**Comments**
Bao et al., 2011, effect of **DBP** on male reproduction in rats 1. Serum sex hormone levels (T, E2, LH, FSH) 2. effects on testes (spermatogenesis, sertoli, testes) 3. Protein expression	0.1-500 mg/kg bw day by gavage, N=5	1. No. E2 and LH 1 data point (CP-5 and CP-6 for E2, less than 3 met for the others) 2. No. Toxicity at >100mg/kg bw) (N/A) 3. No. Vimentin 1 data point (N/A)	1. Early event 2. Apical effect 3. Early event	-	-	1. P_{NMDR} 89 (∩ for LH) 2 Not analysed 3 Not analysed	For T, an ∩ shape trend is observed (reaching 130% of control values) but differences are not statistically significant
Lehmann et al., 2004, effect of **DBP** in utero exposure in male rats 1. Testicular mRNA levels 2. Protein expression 3. Testosterone levels in testes	0.1-500 mg/kg bw day by gavage N=6	1. No (less than 3 met) 2. No, 1 data point 3. No, clearly monotonic	1. Early event 2. Early event 3. Early event	-	-	1. Low for NMDR (MDR for the different mRNA) 2. Not analysed 3. Not analysed	
Andrade et al., 2006a (adult male), effects of **DEHP** in utero and lactation exposure on adult male rats exposed 1. Serum T concentration 2. Sperm morphology, testicular morphometry 3. Sexual behaviour	0.015-405 mg/kg bw day by gavage N=10	1. No (CP5 and CP-6) 2. No (Only 2 CP met) 3. No (Only 2 CP met)	1. Early event 2. Intermediate 3. Apical effect	-	-	Not analysed	1. large within-group variability 3. large within-group variability (SE)

Andrade et al., 2006b (aromatase), effects of **DEHP** in utero and lactation exposure on aromatase activity at PND 1 & 22 in rats 1. Males PND1 2. Females PND1 3. Males PND22 4. Females PND22	0.015-405 mg/kg bw day by gavage N=10	1. Yes U (All checkpoints met) 2. No (All unmet) 3. No (All unmet) 4. No, (Only 3 met)	1. Early event 2. Early event 3. Early event 4. Early event	No	No. (In addition, no effects on apical Repro parameters in #10)	Not analysed	1. large, overlapping SD and plateau at 4 highest doses. Statistics? 4. consistent increase except 1 data point
Grande et al., 2007, effects of **DEHP** in utero and lactation exposure on reproduction in female rats 1. Age vaginal opening 2. Age at 1st estrus 3. Ano-genital distance PND22 4. Number of nipples at PND13	0.015-405 mg/kg bw day by gavage N=10	1. No, increase at high doses 2. No, trend for increase at high doses 3. No effect 4. No effect	1. Intermediate 2. Intermediate 3. Apical 4. Intermediate	-	-	Not analysed	Not included in the Report but added for completeness as reports findings from the same study. Repro parameters not affected (litter size, implantation, birth wt, sex ratio, ano-genital distance at PND22, number of nipples at PND13, …)
Andrade et al., 2006c (juvenile males), effects of **DEHP** in utero and lactation exposure on male offspring in rats 1. Ano-genital distance PND22 2. Number of nipples at PND13 3. Testis weight 4. Tubule diameter 5. Intratesticular testosterone PND1 6. Histopathol. Alterations in testes 7. Age at testis descending 8. Age at preputial separation	0.015-405 mg/kg bw day by gavage N=10	1. No. 2. No. ↑ at 405 only 3. No. ↑ ≥5-135, ↓at 405 4. No effect 5. No effect 6. No. Effects at ≥135mg/kg bw 7. No. No effect 8. No. Trend for delay 9. No	1. Apical 2. Apical 3. Apical 4. Intermediate 5. Early event 6. apical effect 7. Apical 8. Apical 9. Apical	-	-	Not analysed	*1. Increase of one data point of doubtful biological relevance* Not included in the Report but added for completeness as reports findings from the same study. Authors comment*: Body weight at preputial separation was mostly unchanged and significant differences (decreased body weight) were only detected at 0.135, 0.405 and 405 mg/kg/day*

9. Bw at preputial separation							
Christiasen et al., 2010, effects of **DEHP** in utero and lactation exposure on male reproduction in rats 1. Levator ani/bulbocavernosus muscles (LABC) weight 2. Body weight 3. Adrenal weight 4. Number of nipples in male 5. Incidence of male offspring with mild external genital dysgenesis 6. Expression of prostate binding protein subunit C3 (PBPC3) mRNA in ventral prostate 7. Right testis weight 8. Ventral prostate weight 9. Expression of ornithine decarboxylase (ODC) mRNA in ventral prostate 10. Liver weight	3-900 mg/kg bw day by gavage, N=7	1. No (Only 3 met) 2. No (Only 2 met) 3. No (Only 2 met) 4. No, 1 data point (larger ↑at 10 mg/kg) (Only one met) 5. No (Only 1 met) 6. No (Only 1 met) 7.No (Only 1 met) 8. No (Only 1 met) 9.No (All unmet) 10.No (All unmet)	1.-5.; 7.-8.;10. Apical effect 6.;9. early event	-	-	1. Low for NMDR (P_{MDR} 47) 2. Low for NMDR (P_{MDR} 49 MDR) 3. Low for NMDR P_{MDR} (42) 4.Not analysed 5. Not analysed 6. Low for NMDR (P_{MDR} 86) 7. Low for NMDR (P_{MDR} 90) 8. Low for NMDR (P_{MDR} 82) 9. Low for NMDR (P_{MDR} 78) 10. Low for NMDR (P_{MDR} 68.9)	Monotonic effect for ano-genital distance PND1
Do et al 2012, effects of **DEHP** in utero exposure on male reproduction in mice 1. Maternal serum testosterone 2. Fetal male serum testosterone GD18 3. Male offspring testicular testosterone	0.0005-500 mg/kg bw day, feed once daily (GD9-18), N=6	1. Yes U, trend, but probably 1 data point (1µg/kg) (CP-3 and CP-6) 2. No. (Only 3 met) 3. No effect (only 1 met) 4. No. (Only 2 met) 5. No	1. Early event 2. Early event 3. Intermediate 4. Apical 5. Intermediate 6. apical effect	1. ?	1.No? No effect on litter size	1. P_{NMDR} 47 (∩) 2. P_{NMDR} 40.24 (∩) 3. P_{NMDR} 54 (∩) 4. P_{NMDR} 38.5 (∩) 5. P_{NMDR} 31.5 (∩)	2. Increase except 500 or: low control and tox at 500 4. Large SD, 1 data point 6. Decrease at ≥50 mg/kg bw No effects on litter size, birthweight and sex ratio 5 driven mainly by 1 datapoint

4. Ano-genital distance PND18 5. Ratio AGD/BW in males 6. Testis weight		(Only 1 met) 6. No (All unmet)				6. Low for NMDR (P_{MDR} 43.9)	
Grande et al., 2006 (juvenile females) effects of **DEHP** in utero and lactation exposure on female offspring in rats 1. Body weight of offspring at PND1 2. Body weight at vaginal opening 3. Kidney weight of dams 4. Body weight at first estrus	0.015-405 mg/kg bw Gavage, (GD6-PND21), N=11	1. No effects (CP-5 and CP-6) 2. No effects (Only 2 met) 3. No effects (Only one met) 4. No effects (Only 1 met)	1. apical effect 2. apical effect 3. apical effect 4. apical effect	-	-	Not analysed	No effects observed for E2, progesterone or estrus cycling Effects at the heist dose for vaginal and uterine luminal cell height and # of ovarian atretic tertiary follicles
Blystone 2010, rats Multigeneration study, exposure of P0, F1, F2 (3 litters each generation) 1.Testicular malformations 2.Epididymis malformations 3. Pregnancy index (number of females delivering/number of cohabiting pairs) in F3 generation	1.5-10'000mg/kg feed (F3 only up to 7500 mg/kg feed) (0.1-500mg/kg bw day) N=8 (1.5 mg/kg feed in controls)	1.&2.No. Increased incidence at ≥7500 mg/kg feed 3.No. F3 pregnancy index ↓ at 7500mg/kg feed (359 mg/kg bw)F1 at 10000mg/kg feed (543mg/kg bw) did not produce F2	1.-3. apical effect	-	-	Not analysed	In controls, background exposure was measured (Remark*: included to support the Grande/Andrade studies and to highlight the lack of measurement of background exposure in probably all the other studies)*

*CP = checkpoint as defined in the Report (see full list at the "Introduction" section):
CP-3. Can the apparent NMDR be explained by one single potential outlying dose group?
CP-5. Is the steepness of the dose-response curve outside the range of biologically plausible/realistic dose-response shapes?
∏ The symbol U indicates a NMDR with U (or J) shape, the symbol ∩ indicates a NMDR with inverted U (or J) shape
‡ Only addressed when a possible NMDR is confirmed under 1. Presence/ shape of NMDR

A potential NMDR was observed for serum testosterone in Do et al. (2012). In particular, after the exposure of male mice with a wide range of doses of DEHP, the authors reported an inverted U-shaped dose-response curve, characterized on the left side by a monotonic increase of testosterone level (maternal serum testosterone) and a monotonic decrease in the right part of the curve. Andrade et al. (2006a) observed increased serum testosterone levels in male rats exposed at 0.045, 0.405 and 405mgDEHP/kg/day, while levels were similar to control levels at the other doses. However, the mechanistically linked apical effects showed monotonic dose-responses. An additional literature search was conducted for complementing these observations.

Table AII-3 lists all the scientific articles retrieved from the literature search. The NMDRs observed by authors for phthalates have been reported particularly for neuroendocrine, metabolic, and reproductive effects. All the studies were analysed by the WG and the comments are reported in the table.

Table AII-3: Selected publications retrieved from the literature search complemented with the references and citations of the retrieved publications.

Author	NMDR claimed by Authors	Type of study	Comments on non-monotonicity and analysis
Adibi et al. (2010)	Gene expression in the steroidogenesis pathway	Epidemiological study on phthalates' metabolites in placenta	Some indication of NMDR but mostly driven by fluctuations in the 4th quintile. Relevance of gene expression in the steroidogenesis pathway for risk assessment is unclear.
Andrade et al. (2006b)	Brain aromatase activity	DEHP exposure on Wistar rats	Already included in the data set provided by the EFSA External Report. Maybe related to the increase of T leading to overcompensation of the homeostatic feedback mechanism?
Ashley-Martin et al., (2015)	IL-33/TSLP and IgE	Epidemiological study on phthalates' metabolites (first trimester of pregnancy)	Associations were modelled using restricted cubic spline and model fit suggests the presence of NMD for levels of both IL-33/TSLP and IgE. Exposure based on measured levels of MCPP (DBP metabolite)
Barakat et al. (2019)*	Impaired fertility	Environmentally relevant mixture of phthalates (15% DiNP, 21% DEHP, 36% DEP, 15% DBP, 8% DiBP, and 5% BBzP) exposure on CD-1 mice	The lowest dose group (20 µg/kg/day) gave the severest impact for some reproductive endpoints, displaying non-monotonic (gonadal weight at 12 months, StAR and CYP11 expression, sperm concentration) or complex dose response
Binder et al., (2018)	BV (breast total volume)	Epidemiological study on phthalates' metabolites (adolescent girls)	The authors evaluate the dose response (MCNP in urine) by modeling the data using tertiles of exposure. With only three groups limited conclusions of NMDR can be drawn.
Botelho et al. (2009)	Serum cholesterol	DEHP exposure on Wistar rats	Few and too high doses

		N=4 (0, 250, 500, and 750 mg/kg/day) From PND21 to PND51 by gavage	
De Cock et al., (2016)	Birth weight	Epidemiological study on phthalates' metabolites (pregnant women)	The authors evaluate the dose response (MECPP, MEHH in cord plasma) by modeling the data using tertiles of exposure. With only three groups limited conclusions of NMDR can be drawn.
Dai et al. (2015)*	Development of neurotransmitter systems in brain and behavior	DEHP exposure on CD-1 mice	Only 3 doses, all below the NOAEL for reproductive effects, the claim for NMDR cannot be assessed
Do et al. (2012)	Maternal and fetal male serum testosterone level	DEHP exposure on CD-1 mice	Already included in the data set provided by the EFSA External Report.
Du et al., (2018)	Serum Inhibin B (INHB).	Epidemiological study on phthalates' metabolites ('infertile' women)	The observed associations (with MEOHP in urine) appear more inverse and leveling off rather than being non-monotonic.
Gao et al. (2019)	Preterm birth	Epidemiological study on phthalates' metabolites (pregnant women)	No indication of NMDR
Gao et al. (2018)	Neuroendocrine genes in the hypothalamus	DEHP exposure on Sprague–Dawley rats N=4 (0, 2, 10 or 50 mg/kg) From GD14 to 19 by gavage	Few doses but of relevance for the RA
Ge et al. (2007)	Testosterone level, seminal vesicle weight and puberty onset	DEHP exposure on Long-Evans rats N=4 (0, 10, 500, or 750 mg/kg bw/day) PND21 to PND49	Few doses and large range (but doses generally used for tox studies on phthalates); saturation at high doses (general toxicity/MTD?)? "low doses of DEHP (eg, 10 mg/kg body weight) may stimulate androgen production" In the same study also in vitro findings that stress the concept of LC hyperplasia
Hatch et al. (2008)*	Body Mass Index (BMI) in males 12-19 years old	Epidemiological study on phthalates' metabolites	Assessment (MEHHP in urine) based on quartiles, lack of consistency as different shapes are observed for other ages and females and for quartiles for other phthalates metabolites
Hatcher et al. (2019)*	Neuroendocrine genes in the amygdala	DEHP exposure on CD-1 mice	NMDR U shape observed for Esr1 and Nr3c2; inverted-U shape observed for Drd2 and Esr2. Individual measurements provided supporting the assessment

Hu et al. (2020)	Preterm birth	Epidemiological study on phthalates (add phthalates covered by the NMDR)(first trimester of pregnancy)	Assessment (seven different metabolites in urine) based on quartiles, Visual inspection appears to suggests NMDR for some phthalate metabolites (e.g. MCPP). Main risk factors for preterm births include infections, high blood pressure and diabetes but in many cases the causes are unknown. In that perspective the biological explanation for the apparent NMDR in this study is unclear
Huang et al. (2019)	Lipid metabolism	DEHP and DINP exposure on Kunming mice N=3 (0.048 or 4.8 mg/kg) PND0 to 21	Number of doses not adequate for determining NDMR but of relevance
James-Todd et al. (2012)	Diabetes	Epidemiological study on phthalates(women)	The suspected NMDR (MnBP and ∑DEHP metabolites in urine) is driven by the 3rd quartile. The role of chance finding by some sort of formal testing or modelling is not evaluated.
Kasper-Sonnenberg, et al., (2017)	Pubertal development	Epidemiological study on phthalates (children)	NMDR claims are proposed due to non-linear associations (MEHP and cx-MEPP in urine), but data not presented.
Lee et al. (2004)	Pituitary weight and endocrine alterations	DBP exposure on Sprague Dawley rats N=5 (0, 20, 200, 2000 and 10,000 ppm) GD15 to PND21 by diet	N=5 but MTD
Lind and Lind (2011)	Atherosclerotic plaques	Epidemiological study on phthalates	The suspected NMDR (MMP in serum) is driven by one of the quintiles. The role of chance finding by some sort of formal testing or modelling is not evaluated.
Majeed et al. (2017)	Blood serum parameters (cholesterol, glucose and LDH)	DBP exposure on albino rats N=3 (0, 10, 50 mg/kg/bw) For 13 weeks, by diet	Not enough doses to establish a NMDR but "*Further low-dose investigations are needed to assess non-monotonic dose responses*."
Meeker and Ferguson (2011)	Free total triiodothyronine	Epidemiological study on phthalates' metabolites	Although the dose response between MEHHP and free T3 may appear non-monotonic an alternative explanation is that the decrease in free T3 is simply levelling off.

Meeker et al. (2009)*	Testosterone	Epidemiological study on phthalates	Slight increase in T serum levels at the 2nd quintile and clear reduction at the 5th quintile
Oudir et al. (2018)	Serum testosterone level	DEHP exposure on Wistar rats N=4 (0, 0.5, 50, 5000 µg/kg bw/day) From PND 21 to 120, by gavage	Retrieved also from the first literature search Few doses but of relevance
Pan et al., (2011)		Epidemiological study on phthalates' metabolites (Workers)	Results indicate the activation of the feedback mechanism for keeping T levels also at exposure levels well below US HBGVs
Philippat et al. (2012)	Birth weight and birth length	Epidemiological study on phthalates' metabolites (pregnant women)	No indication for NDR for these outcomes
Pocar et al. (2012)*	Reproductive endpoints (testis and ovary weight, cleavage rate, blastocyst rate,	DEHP exposure on CD-1 mice	Only two doses Additional information on dysregulation of HPG feedback
Repouskou et al. (2019)	AGD, histopathological changes, hormone levels, steroidogenesis and gonad aromatase	Phthalate mixture exposure on C57/BL6 mice N=4 (0, 0.26, 2.6 and 13 mg/kg/d) Gestational exposure (From GD 0.5), by diet	Few doses but of relevance
Stroustrup et al., (2018)	Beneficial?	Epidemiological study on phthalates' metabolites (very low birth weight infants)	The presence of NMDR is not evaluated and no data to evaluate by visual inspection or other means are reported
Wang et al. (2016a)*	Behavioral effects	DEHP exposure on ICR mice	Inverted-U shape for social play and investigation times in pubertal males, lead by a single dose (50 mg/kg/day)
Wang et al. (2016b)*	Body weight, and hormone receptors	DEHP exposure on ICR mice females	Inverted-U shape for body weight lead by a single dose (1 mg/kg/day) NMDR U shape for estrogen receptor and phosphorylation of ERK1/2
Wang et al. (2018)	T3 or the T3/T4 ratio	Epidemiological study on phthalates' metabolites (workers)	There are some indication of NMDR (the model fit based on restricted cubic spline regression confirms that).

* Study retrieved in the complementary search on references + citations of retrieved publications.

The literature search resulted in the identification of Oudir et al. (2018), where the authors observed a dose-dependent increase of testosterone accompanied by the hyperplasia of Leydig cells after the exposure of low doses of DEHP, followed by decrease at very high doses.

Plausibility assessment for a NMDR for testosterone levels

As the effects on testosterone levels have been identified by the CEP Panel and others as a critical step in the pathway for the reproductive effects, the NMDR assessment has focused on this endpoint. As aromatase is involved in the metabolism of testosterone to oestradiol there is also a connection with the NMDR observed for aromatase inhibition in DEHP exposed animals.

Table AII-4 summarizes the experimental studies providing information to assess the significance and biological plausibility of the NMDR observed for testosterone levels in response to phthalates' exposure. Only few studies provide a complete representation of this (non-monotonic) dose-response curve, but many other studies using fewer doses provide complementary information supporting an increase in testosterone levels at low doses of phthalates, which are opposite to the confirmed reduction at high doses linked to the reproductive effects by the CEP Panel and others (e.g. NAS 2017 meta-analysis).

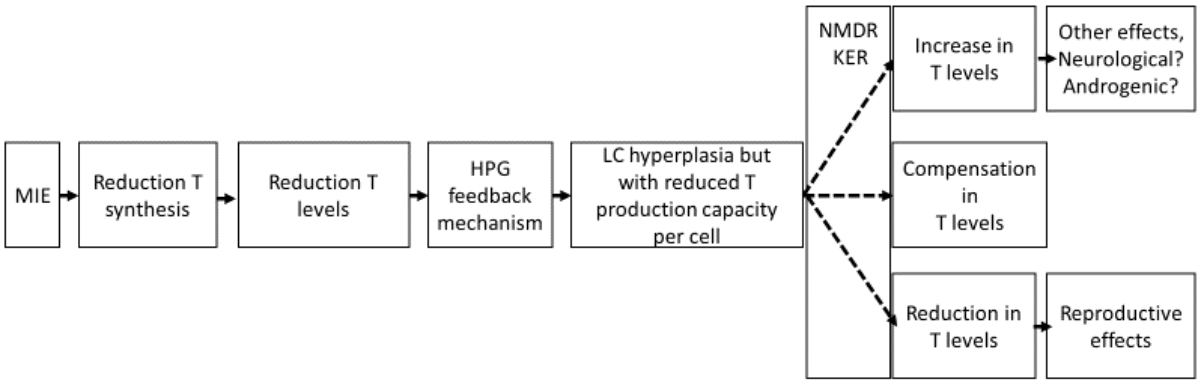


Figure AII-1: Hypothesis of an AOP-based mechanistic understanding of the inverted U-shaped curve for testosterone level as early/intermediate event. T (testosterone), LC (Leydig cells), MIE (molecular initiating event), KER (key event relationship). The three arrows in the KER box represents options for the net result of the combination or an initial reduction and the feedback mechanisms, the NMDR with increase of T levels at low doses is linked to the overcompensation of the homeostatic response.

The non-monotonic dose-response for testosterone level has been also proposed by Ge et al. (2007a) supported by *in vivo* evidence in rats after DEHP exposure, and *in vitro* after MEHP exposure. In this study the increase in testosterone levels was associated with the advancement of puberty onset, contrarily to the opposite effect observed at high doses. Moreover, the same group of researchers observed the androgen stimulation associated with phthalates-induced Leydig cells aggregation and hyperplasia (Akingbemi et al. (2004); (Lin et al., 2008).

The low-dose-induced plasma testosterone increase has been reported after oral chronic administration as well as after subacute inhalation of DEHP in rats (Kurahashi et al. (2005)). Besides, the same effect was noticed in different species, rodents (i.e. rats and mice) and non-rodents as observed by Ljungvall et al. (2005) where boars exposed to DEHP during the prepubertal period showed an increase of testosterone level with a concomitant increase of Leydig cell area 4.5 months after the exposure period.

The mechanism of action hypothesized by the authors is often the disturbance in the HPG feedback mechanism. Studies such as Akingbemi et al. (2004) provide evidence compatible with the role of the feedback mechanism in the NMDR, as DEHP exposure triggers a reduction in basal and LH-induced testosterone production, LH increase, and Leydig cells hyperplasia that under certain conditions produces an increase in testosterone levels. The non-monotonic dose-response on rat brain aromatase activity observed by Andrade et al. (2006b) seems to be in concordance with this hypothesis. A possible relationship between phthalates exposure and aromatase suppression is suggested also by Meeker et al. (2009), and *in vitro* and *in silico* molecular docking studies confirm the elevated binding affinity of phthalates to CYP19A1 (Gupta et al. (2010), Ahmad et al. (2017)).

There is information covering experimental, human and epidemiological studies associating testosterone increase with neurological and neurodevelopmental effects (Qi et al. (2018), Nakano et al. (2010), Hines (2003), Schwarz et al. (2011)). Other authors link testosterone stimulation with apical effects associated to "over-masculinization", as expected for overexpression of androgens (Hotchkiss et al. (2007), Martin et al. (1998)). Accordingly, some epidemiological studies link phthalate exposure with compatible metabolomic alterations (Zhou et al. (2018), neurodevelopmental (Braun (2017), Engel et al. (2018)) and effects attributable to hyperandrogenism (Colon et al. (2000)). In humans, brain aromatase has been associated to personality traits and neurobehavioral disorders (Takahashi et al. (2018), Sarachana et al. (2011)), providing an additional mechanistic link.

In line with a weight of evidence approach, all this data provide relevant insight supporting the biological relevance of the NMDR observed by Do et al. (2012). This information may also serve as building blocks for an AOP-based mechanistic understanding of the inverted U-shaped curve for testosterone level as early/intermediate event. The elements linked to the NMDR assessment are described in Figure AII-1, the phthalates-induced reduction in testosterone (T) synthesis leads to an initial reduction in T levels that triggers the compensatory feedback mechanism which includes Leydig cell hyperplasia. A plausible hypothesis for the NMDR is that the non-monotonicity is associated to the key event relationship (KER); the continuous stimulus of the feedback mechanisms results in Leydig cells hyperplasia but with reduced T production capacity. The combination of both processes could explain the non-monotonic response in T levels, at low doses the increase in cell number not only compensate but exceeds the reduction in the production capacity, resulting in overall T increase. Under certain conditions a compensation is achieved, and T levels remain unchanged. At high doses the cellular increase is insufficient and a net reduction in T levels is observed.

This is just one of the possible pathways explaining *in vivo* NMDR in T response connected to Leydig cell hyperplasia. Another possibility could be that the NMDR on testosterone production observed *in vitro* for the main metabolite MEHP, under certain circumstances could be directly responsible for the *in vivo* NMDR, particularly in those cases that DEHP exposure is not linked with Leydig cell hyperplasia. This pathway could be complementary or alternative to the AOP proposal considering the complexity and differences in responses at different critical target windows. A full assessment of the MIEs and early steps in the pathways is outside the scope of this assessment. It should be also noted that the changes in the steroidogenesis induced by DEHP are linked to the window and duration of the exposure, and that the roles of the feed-back control of testosterone production and DEHP effects on aromatase have been previously postulated (Ljungvall et al., 2005).

The consequences of testosterone reduction in specific windows are related to the observed reproductive effects. There are some indications associating increased testosterone levels with non-reproductive apical effects similar or associated to those described as no fully confirmed but of possible concern in the CEP opinion. The establishment of a quantitative association with the related adverse apical outcomes would require a full assessment and is outside this mandate. In addition, the etiology of the hypothesized effects is mostly multifactorial and still poorly understood. The extent of the effect is likely to be dependent on the specific exposure window, varying among sexes and individuals. The final effect may also depend on the effective internal concentration of MEHP, which seems to be partly responsible for the effect of DEHP.

Author	Summary of the experimental studies supporting the phthalates' NMDR for testosterone level	
	DEHP	**MEHP**
Do et al. 2012	Pregnant CD-1 mice fed with 0.5–500,000 µg/kg/day DEHP from GD 9–18 Inverted U-shape dose response curve noticed from 0-500,000 µg/kg/day for maternal serum testosterone in the dam (n=9-20/group) and fetal male serum testosterone in the pups (n=11-21/group)	Serum concentration of MEHP increase in a monotonic manner
Andrade et al. 2006	Female Wistar rats treated with 0.015-405 mg/kg/day DEHP from GD 6 to LD 21 by gavage Increase of serum T noticed from 0.045-0.405 and at 405 in male adult offspring (n=19–20 per dose)	
Oudir et al. 2018	Male Wistar rats treated with 0.5-500 µg/kg bw/day from PND 21 to 120 (during pre-pubertal, pubertal and post-pubertal period) by gavage Increase of serum T noticed at 0.5 with con concomitant LC hyperplasia in male rats (n=10/group)	
Zhao et al. 2012		Long-Evans rats used for the isolation of LC at PND 21, 35 and 49. LC were exposed ex vivo to 2-2000 µM of MEHP Increase of testosterone production noticed at 20–200 µM in adult Leydig cells (ALC) (PND49) Decrease of testosterone production noticed at 2000 µM in ALC (PND49)
Jones et al. 2015		Sprague Dawley rats used for the isolation of testes at PND3. The organ was cultured ex vivo and exposed to 10 µM Stimulatory effect on basal testosterone production that was normalized by GEN
Kurahashi et al. 2015	Prepubertal male Wistar Rats exposed to 5 or 25 mg/m^3 of DEHP 6h/day, from PND28 to 56 or 84, by inhalation. Increase plasma testosterone concentration (n=12/group)	
Ge et al. 2007	Long-Evans male rats (n= 10/group) treated with DEHP (10, 500, or 750 mg/kg) from PND 21 to 48, by gavage Biphasic effect on testosterone level with ↑ at 10 and ↓ at 750 with biphasic effect on onset of puberty at the same doses	MEHP exposure (10^{-9} - 10^{-2} M) in vitro on LC isolated from Long-Evans male rats at PND35 Biphasic effect on LH-induced testosterone production with ↑ at 100 µM and ↓ >10 mM.

Akingbemi et al. 2004	Long-Evans rats (n= 10/group) treated with DEHP (0, 10, or 100 mg/kg/day) from PND 21 to 48, 90, or 120 by gavage. Increase of serum T level with concomitant elevated serum LH, E2 and LC hyperplasia (confirmed by increase cell cycle proteins).	Serum concentration of MEHP increase in a monotonic manner
Lin et al. 2008	Pregnant Long–Evans rats female were treated from GD2 to GD20 (n>=6/group) with DEHP (0, 10, 100, or 750 mg/kg/day) by gavage. Effects examined at GD21 in foetus. Biphasic effect on testicular testosterone (+50% at 10; -66% at 750) concordant with the biphasic effect on IGF1 and KITL gene expression (↑ at 10). Effect on fetal Leydig cells aggregation.	
Gunnarsson et al. 2008		MEHP exposure (25-100 µM) on mouse Leydig tumor cell line (MLTC-1) and on granulosa tumor cell line (KK-1). - Stimulatory effect on testosterone as well as progesterone (monotonic increase from 25 to 100 µM)
Savchuk et al. 2015		MEHP ex vivo exposure (1, 3, 10, 30, and 90 µM) on cells from C57BL/6j and CBA/Lac mouse Stimulation of basal steroidogenesis at 90 with concomitant with upregulation of StAR protein expression, ATP depletion and increase SOD generation, but increase viability.
Forgacs et al. 2012		MEHP ex vivo exposure (3,10,30,100, 300 µM) on BLTK1 Murine Leydig Cells Increase basal testosterone level at 100-300
Ljungvall et al. 2005	Boar (n=4/group) exposed to of DEHP (50mg/kg) twice a week for 5 weeks (prepubertal exposure) by i.m. injection Increase T 4.5 months after exposure with concomitant increase of the LC area	

Answer to the questions (proposed approach)

What is the experimental evidence for the effect observed (*in vitro* / *in vivo*? Other?)

• One experimental study, supported by the statistical evaluation, showing a non-monotonic dose-response curve (Do et al. 2012). *In vitro* studies showing also an inverted U-shaped curve (Zhao et al. (2012), Ge et al. (2007a)).
• Significant amount of complementary data, e.g. covering part of the dose-response curve, supporting these findings, such as:
  ○ Well-known monotonic decrease in serum testosterone induced by high doses of phthalates (clearly showed by the NAS meta-analysis).
  ○ Several studies *in vitro*, *ex vivo* and *in vivo*, covering different species and routes of exposure, reporting increase of testosterone after phthalates-low dose exposure (0.001-0.5 mg/kg bw *in vivo*; 20–200 µM *in vitro*) .

What is the biological relevance of the effects observed? Can a (quantitative) relation between the observed effect and an adverse outcome be established? Ideally: Could a mechanistic sequence (AOP) be partially or fully established? If yes, specific considerations need to be applied and a diversion from the current methodologies for RA may be needed
• There is information supporting that the NMDR observed for the intermediate effect could be related to the disturbance of the HPG feedback mechanism. A possible mechanistic interpretation is that the combination of two different phenomena, a) the phthalate's induced reduction in T production capacity by Leydig cells, and b) the compensatory Leydig cells hyperplasia triggered by the feedback mechanism, could result in a NM key event relationship and NMDR for T levels. • The connection between the reduction in T levels during critical windows and reproductive effects is well established • The adverse reproductive effects are mechanistically linked to the reduction of T levels during a critical period. The net increase in T levels, plausibly linked to the overcompensation of the feedback mechanisms, will not result in these reproductive effects, but may be connected to different adverse outcomes. The critical period for adverse outcomes and the critical exposure window for T increases, may be different from those related to the decrease in T levels. • There are experimental and epidemiological studies that link the increase of testosterone with neurodevelopmental effect and other effects linked to over-masculinization (as expected for overexpression of androgens) • There also studies associating phthalate exposure with similar effects • The establishment of a quantitative relationship between the increase in T levels and the observed effects should consider the multifactorial etiology of the referred adverse outcomes that the effect is likely to be dependent on the specific exposure window, varying among sexes and individuals; and requires a full assessment of the information on phthalates and non-reproductive effects which is outside the scope of this mandate
If information is lacking on whether an observed effect can lead to an adverse outcome, additional testing may be needed. Here NAMs would be of relevance given the need for identifying a mechanistic sequence of events.

- The possible hypothesis to be explored is that phthalates may have a NMDR for the intermediate event testosterone level linked to different routes to adverse effects:
    - High phthalate exposure produces anti-androgenic effects linked to testosterone decrease and the associated reproductive adverse outcomes (mostly malformations);
    - Low phthalate exposure levels may be linked to testosterone increase (e.g. postulated through the overstimulation of the compensatory feedback mechanism) and may be associated with different kinds of adverse outcomes, signalling to masculinization in females (i.e. hirsutisms) and neurodevelopmental effects associated to testosterone increase.
- This assessment is outside the scope of this mandate.

References

Adibi JJ, Whyatt RM, Hauser R, Bhat HK, Davis BJ, Calafat AM, Hoepner LA, Perera FP, Tang D and Williams PL, 2010. Transcriptional Biomarkers of Steroidogenesis and Trophoblast Differentiation in the Placenta in Relation to Prenatal Phthalate Exposure. Environmental Health Perspectives, 118:291-296. doi:10.1289/ehp.0900788

Ahmad S, Khan MF, Parvez S, Akhtar M and Raisuddin S, 2017. Molecular docking reveals the potential of phthalate esters to inhibit the enzymes of the glucocorticoid biosynthesis pathway. Journal of Applied Toxicology, 37:265-277. doi:10.1002/jat.3355

Akingbemi BT, Ge RS, Klinefelter GR, Zirkin BR and Hardy MP, 2004. Phthalate-induced Leydig cell hyperplasia is associated with multiple endocrine disturbances. Proceedings of the National Academy of Sciences of the United States of America, 101:775-780. doi:10.1073/pnas.0305977101

Andrade AJM, Grande SW, Talsness CE, Gericke C, Grote K, Golombiewski A, Sterner-Kock A and Chahoud I, 2006a. A dose response study following in utero and lactational exposure to di-(2-ethylhexyl) phthalate (DEHP): Reproductive effects on adult male offspring rats. Toxicology, 228:85-97. doi:10.1016/j.tox.2006.08.020

Andrade AJM, Grande SW, Talsness CE, Grote K and Chahoud I, 2006b. A dose-response study following in utero and lactational exposure to di-(2-ethylhexyl)-phthalate (DEHP): Non-monotonic dose-response and low dose effects on rat brain aromatase activity. Toxicology, 227:185-192. doi:10.1016/j.tox.2006.07.022

Andrade AJ, Grande SW, Talsness CE, Grote K, Golombiewski A, Sterner-Kock A and Chahoud I, 2006c. A dose-response study following in utero and lactational exposure to di-(2-ethylhexyl) phthalate (DEHP): effects on androgenic status, developmental landmarks and testicular histology in male offspring rats. Toxicology, 225:64-74. doi: 10.1016/j.tox.2006.05.007

Arukwe A, Ibor OR and Adeogun AO, 2017. Biphasic modulation of neuro- and interrenal steroidogenesis in juvenile African sharptooth catfish (Clarias gariepinus) exposed to waterborne di-(2-ethylhexyl) phthalate. General and Comparative Endocrinology, 254:22-37. doi:10.1016/j.ygcen.2017.09.007

Ashley-Martin J, Dodds L, Levy AR, Platt RW, Marshall JS and Arbuckle TE, 2015. Prenatal exposure to phthalates, bisphenol A and perfluoroalkyl substances and cord blood levels of IgE, TSLP and IL-33. Environmental research, 140:360-368. doi: 10.1016/j.envres.2015.04.010

Bao A-M, Man X-M, Guo X-J, Dong H-B, Wang F-Q, Sun H, Wang Y-B, Zhou Z-M and Sha J-H, 2011. Effects of di-n-butyl phthalate on male rat reproduction following pubertal exposure. Asian journal of andrology, 13:702-709. doi: 10.1038/aja.2011.76

Barakat R, Seymore T, Lin PP, Park CJ and Ko CJ, 2019. Prenatal exposure to an environmentally relevant phthalate mixture disrupts testicular steroidogenesis in adult male mice. Environ Res, 172:194-201. doi:10.1016/j.envres.2019.02.017

Beausoleil C, Beronius A, Bodin L, Bokkers BGH, Boon PE, Burger M, Cao Y, De Wit L, Fischer A, Hanberg A, Leander K, Litens-Karlsson S, Rousselle C, Slob W, Varret C, Wolterink G and Zilliacus J, 2016. Review of non-monotonic dose-responses of substances for human risk assessment. EFSA Supporting Publications, 13:1027E. doi: 10.2903/sp.efsa.2016.EN-1027

Bello UM, Madekurozwa MC, Groenewald HB, Aire TA and Arukwe A, 2014. The effects on steroidogenesis and histopathology of adult male Japanese quails (Coturnix coturnix japonica) testis following pre-pubertal exposure to di(n-butyl) phthalate (DBP). Comparative Biochemistry and Physiology C-Toxicology & Pharmacology, 166:24-33. doi:10.1016/j.cbpc.2014.06.005

Binder AM, Corvalan C, Pereira A, Calafat AM, Ye XY, Shepherd J and Michels KB, 2018. Prepubertal and Pubertal Endocrine-Disrupting Chemical Exposure and Breast Density among Chilean Adolescents. Cancer Epidemiology Biomarkers & Prevention, 27:1491-1499. doi: 10.1158/1055-9965.Epi-17-0813

Blystone CR, Kissling GE, Bishop JB, Chapin RE, Wolfe GW and Foster PMD, 2010. Determination of the Di-(2-Ethylhexyl) Phthalate NOAEL for Reproductive Development in the Rat: Importance of the Retention of Extra Animals to Adulthood. Toxicological Sciences, 116:640-646. doi: 10.1093/toxsci/kfq147

Botelho GGK, Golin M, Bufalo AC, Morais RN, Dalsenter PR and Martino-Andrade AJ, 2009. Reproductive Effects of Di(2-ethylhexyl)phthalate in Immature Male Rats and Its Relation to Cholesterol, Testosterone, and Thyroxin Levels. Archives of Environmental Contamination and Toxicology, 57:777-784. doi:10.1007/s00244-009-9317-8

Braun JM, 2017. Early-life exposure to EDCs: role in childhood obesity and neurodevelopment. Nat Rev Endocrinol, 13:161-173. doi:10.1038/nrendo.2016.186

Chen X, Liu YN, Zhou QH, Leng L, Chang Y and Tang NJ, 2013a. Effects of Low Concentrations of Di-(2-ethylhexyl) and Mono-(2-ethylhexyl) Phthalate on Steroidogenesis Pathways and Apoptosis in the Murine Leydig Tumor Cell Line MLTC-1. Biomedical and Environmental Sciences, 26:986-989. doi:10.3967/bes2013.034

Chen X, Zhou QH, Leng L, Chen X, Sun ZR and Tang NJ, 2013b. Effects of di(n-butyl) and monobutyl phthalate on steroidogenesis pathways in the murine Leydig tumor cell line MLTC-1. Environmental Toxicology and Pharmacology, 36:332-338. doi:10.1016/j.etap.2013.04.013

Cheng LJ and Cheng TS, 2012. Oxidative effects and metabolic changes following exposure of greater duckweed (Spirodela polyrhiza) to diethyl phthalate. Aquatic Toxicology, 109:166-175. doi:10.1016/j.aquatox.2011.10.003

Christiansen S, Boberg J, Axelstad M, Dalgaard M, Vinggaard AM, Metzdorff SB and Hass U, 2010. Low-dose perinatal exposure to di(2-ethylhexyl) phthalate induces anti-androgenic effects in male rats. Reprod Toxicol, 30:313-321. doi: 10.1016/j.reprotox.2010.04.005

Colon I, Caro D, Bourdony CJ and Rosario O, 2000. Identification of phthalate esters in the serum of young Puerto Rican girls with premature breast development. Environmental Health Perspectives, 108:895-900. doi:10.2307/3434999

Dai YH, Yang YL, Xu XH and Hu YZ, 2015. Effects of uterine and lactational exposure to di-(2-ethylhexyl) phthalate on spatial memory and NMDA receptor of hippocampus in mice. Hormones and Behavior, 71:41-48. doi:10.1016/j.yhbeh.2015.03.008

de Cock M, De Boer MR, Lamoree M, Legler J and Van de Bor M, 2016. Prenatal exposure to endocrine disrupting chemicals and birth weight-A prospective cohort study. Journal of Environmental Science and Health Part a-Toxic/Hazardous Substances & Environmental Engineering, 51:178-185. doi: 10.1080/10934529.2015.1087753

Dikmen BY, Alpay M, Kismali G, Filazi A, Kuzukiran O and Sireli UT, 2015. In Vitro Effects of Phthalate Mixtures on Colorectal Adenocarcinoma Cell Lines. Journal of Environmental Pathology Toxicology and Oncology, 34:115-123. doi:10.1615/JEnvironPatholToxicolOncol.2015013256

Do RP, Stahlhut RW, Ponzi D, Vom Saal FS and Taylor JA, 2012. Non-monotonic dose effects of in utero exposure to di(2-ethylhexyl) phthalate (DEHP) on testicular and serum testosterone and anogenital distance in male mouse fetuses. Reprod Toxicol, 34:614-621. doi:10.1016/j.reprotox.2012.09.006

Du YY, Guo N, Wang YX, Hua X, Deng TR, Teng XM, Yao YC and Li YF, 2018. Urinary phthalate metabolites in relation to serum anti-Mullerian hormone and inhibin B levels among women from a fertility center: a retrospective analysis. Reproductive Health, 15. doi: 10.1186/s12978-018-0469-8

ECHA, 2017a. Committee for Risk Assessment (RAC) and Committee for Socio-economic Analysis (SEAC). Opinion on an Annex XV dossier proposing restrictions on four phthalates (DEHP, BBP, DBP, DIBP).

ECHA, 2017b. Committee for Risk Assessment (RAC) and Committee for Socio-economic Analysis (SEAC). Background document to the Opinion on the Annex XV dossier proposing restrictions on four phthalates (DEHP, BBP, DBP, DIBP)

EFSA CEP Panel (EFSA Panel on Food Contact Materials, Enzymes and Processing Aids), Silano V, Barat Baviera JM, Bolognesi C, Chesson A, Cocconcelli PS, Crebelli R, Gott DM, Grob K, Lampi E, Mortensen A, Riviere G, Steffensen I-L, Tlustos C, Van Loveren H, Vernis L, Zorn H, Cravedi J-P, Fortes C, Tavares Pocas MF, Waalkens-Berendsen I, Wolfle D, Arcella D, Cascio C, Castoldi AF, Volk K and Castle L, 2019. Scientific Opinion on the update of the risk assessment of di-butylphthalate (DBP), butyl-benzyl-phthalate (BBP), bis(2 ethylhexyl)phthalate (DEHP), di-isononylphthalate (DINP) and di-isodecylphthalate (DIDP) for use in food contact materials. EFSA Journal 2019;17(12):5838, 85 pp. https://doi.org/10.2903/j.efsa.2019.5838

Engel SM, Villanger GD, Nethery RC, Thomsen C, Sakhi AK, Drover SSM, Hoppin JA, Zeiner P, Knudsen GP, Reichborn-Kjennerud T, Herring AH and Aase H, 2018. Prenatal Phthalates, Maternal Thyroid Function, and Risk of Attention-Deficit Hyperactivity Disorder in the Norwegian Mother and Child Cohort. Environmental Health Perspectives, 126:11. doi:10.1289/ehp2358

Fisher BG, Frederiksen H, Andersson AM, Juul A, Thankamony A, Ong KK, Dunger DB, Hughes IA and Acerini CL, 2018. Serum Phthalate and Triclosan Levels Have Opposing Associations With Risk Factors for Gestational Diabetes Mellitus. Frontiers in Endocrinology, 9. doi:10.3389/fendo.2018.00099

Gao N, Hu R, Huang Y, Dao L, Zhang C, Liu Y, Wu L, Wang X, Yin W, Gore AC and Sun Z, 2018. Specific effects of prenatal DEHP exposure on neuroendocrine gene expression in the developing hypothalamus of male rats. Arch Toxicol, 92:501-512. doi:10.1007/s00204-017-2049-z

Gao H, Wang YF, Huang K, Han Y, Zhu YD, Zhang QF, Xiang HY, Qi J, Feng LL, Zhu P, Hao JH, Tao XG and Tao FB, 2019. Prenatal phthalate exposure in relation to gestational age and preterm birth in a prospective cohort study. Environmental research, 176. doi: 10.1016/j.envres.2019.108530

Ge RS, Chen GR, Dong Q, Akingbemi B, Sottas CM, Santos M, Sealfon SC, Bernard DJ and Hardy MP, 2007a. Biphasic effects of postnatal exposure to diethylhexylphthalate on the timing of puberty in male rats. Journal of Andrology, 28:513-520. doi:10.2164/jandrol.106.001909

Ge RS, Chen GR, Tanrikut C and Hardy MP, 2007b. Phthalate ester toxicity in Leydig cells: Developmental timing and dosage considerations. Reproductive Toxicology, 23:366-373. doi:10.1016/j.reprotox.2006.12.006

Grande SW, Andrade AJM, Talsness CE, Grote K and Chahoud I, 2006. A Dose-Response Study Following In Utero and Lactational Exposure to Di(2-ethylhexyl)phthalate: Effects on Female Rat Reproductive Development. Toxicological Sciences, 91:247-254. doi: 10.1093/toxsci/kfj128

Grande SW, Andrade AJ, Talsness CE, Grote K, Golombiewski A, Sterner-Kock A and Chahoud I, 2007. A dose-response study following in utero and lactational exposure to di-(2-ethylhexyl) phthalate (DEHP): reproductive effects on adult female offspring rats. Toxicology, 229:114-122. doi: 10.1016/j.tox.2006.10.005

Gupta RK, Singh JM, Leslie TC, Meachum S, Flaws JA and Yao HHC, 2010. Di-(2-ethylhexyl) phthalate and mono- (2-ethylhexyl) phthalate inhibit growth and reduce estradiol levels of antral follicles in vitro. Toxicology and Applied Pharmacology, 242:224-230. doi:10.1016/j.taap.2009.10.011

Hannon PR, Brannick KE, Wang W and Flaws JA, 2015. Mono(2-Ethylhexyl) Phthalate Accelerates Early Folliculogenesis and Inhibits Steroidogenesis in Cultured Mouse Whole Ovaries and Antral Follicles. Biology of Reproduction, 92. doi:10.1095/biolreprod.115.129148

Hatch EE, Nelson JW, Qureshi MM, Weinberg J, Moore LL, Singer M and Webster TF, 2008. Association of urinary phthalate metabolite concentrations with body mass index and waist circumference: a cross-sectional study of NHANES data, 1999-2002. Environmental Health, 7. doi:10.1186/1476-069x-7-27

Hatcher KM, Willing J, Chiang C, Rattan S, Flaws JA and Mahoney MM, 2019. Exposure to di-(2-ethylhexyl) phthalate transgenerationally alters anxiety-like behavior and amygdala gene expression in adult male and female mice. Physiol Behav, 207:7-14. doi:10.1016/j.physbeh.2019.04.018

Hines M, 2003. Sex steroids and human behavior: Prenatal androgen exposure and sex-typical play behavior in children. In: Panzica G and Melcangi RC (eds.). Steroids and the Nervous System. New York, New York Acad Sciences. pp. 272-282.

Hotchkiss AK, Lambright CS, Ostby JS, Parks-Saldutti L, Vandenbergh JG and Gray LE, 2007. Prenatal testosterone exposure permanently masculinizes anogenital distance, nipple development, and reproductive tract morphology in female Sprague-Dawley rats. Toxicological Sciences, 96:335-345. doi:10.1093/toxsci/kfm002

1418 Hu JMY, Arbuckle TE, Janssen P, Lanphear BP, Braun JM, Platt RW, Chen A, Fraser WD and
1419 McCandless LC, 2020. Associations of prenatal urinary phthalate exposure with preterm
1420 birth: the Maternal-Infant Research on Environmental Chemicals (MIREC) Study. Canadian
1421 Journal of Public Health, 111:333-341. doi: 10.17269/s41997-020-00322-5

1422 Huang YC, Sun FJ, Tan HL, Deng YF, Sun ZQ, Chen HX, Li J and Chen D, 2019. DEHP and
1423 DINP Induce Tissue- and Gender-Specific Disturbances in Fatty Acid and Lipidomic Profiles
1424 in Neonatal Mice: A Comparative Study. Environmental Science & Technology, 53:12812-
1425 12822. doi:10.1021/acs.est.9b04369

1426 James-Todd T, Stahlhut R, Meeker JD, Powell SG, Hauser R, Huang TY and Rich-Edwards J,
1427 2012. Urinary Phthalate Metabolite Concentrations and Diabetes among Women in the
1428 National Health and Nutrition Examination Survey (NHANES) 2001-2008. Environmental
1429 Health Perspectives, 120:1307-1313. doi:10.1289/ehp.1104717

1430 Jones S, Boisvert A, Francois S, Zhang L and Culty M, 2015. In utero exposure to di-(2-
1431 ethylhexyl) phthalate induces testicular effects in neonatal rats that are antagonized by
1432 genistein cotreatment. Biol Reprod, 93:92. doi:10.1095/biolreprod.115.129098

1433 Kasper-Sonnenberg M, Wittsiepe J, Wald K, Koch HM and Wilhelm M, 2017. Pre-pubertal
1434 exposure with phthalates and bisphenol A and pubertal development. Plos One, 12. doi:
1435 10.1371/journal.pone.0187922

1436 Kurahashi N, Kondo T, Omura M, Umemura T, Ma MY and Kishi R, 2005. The effects of
1437 subacute inhalation of Di (2-ethylhexyl) phthalate (DEHP) on the testes of prepubertal
1438 Wistar rats. Journal of Occupational Health, 47:437-444. doi:10.1539/joh.47.437

1439 Lee KY, Shibutani M, Takagi H, Kato N, Takigami S, Uneyama C and Hirose M, 2004. Diverse
1440 developmental toxicity of di-n-butyl phthalate in both sexes of rat offspring after maternal
1441 exposure during the period from late gestation through lactation. Toxicology, 203:221-238.
1442 doi:10.1016/j.tox.2004.06.013

1443 Lehmann KP, Phillips S, Sar M, Foster PM and Gaido KW, 2004. Dose-dependent alterations in
1444 gene expression and testosterone synthesis in the fetal testes of male rats exposed to di
1445 (n-butyl) phthalate. Toxicol Sci, 81:60-68. doi: 10.1093/toxsci/kfh169

1446 Lin H, Ge RS, Chen GR, Hu GX, Dong L, Lian QQ, Hardy DO, Sottas CM, Li XK and Hardy MP,
1447 2008. Involvement of testicular growth factors in fetal Leydig cell aggregation after
1448 exposure to phthalate in utero. Proceedings of the National Academy of Sciences of the
1449 United States of America, 105:7218-7222. doi:10.1073/pnas.0709260105

1450 Lind PM and Lind L, 2011. Circulating levels of bisphenol A and phthalates are related to carotid
1451 atherosclerosis in the elderly. Atherosclerosis, 218:207-213.
1452 doi:10.1016/j.atherosclerosis.2011.05.001

1453 Ljungvall K, Karlsson P, Hulten F, Madej A, Norrgren L, Einarsson S, Rodriguez-Martinez H and
1454 Magnusson U, 2005. Delayed effects on plasma concentration of testosterone and testicular
1455 morphology by intramuscular low-dose di(2-ethylhexyl)phthalate or oestradiol benzoate in
1456 the prepubertal boar. Theriogenology, 64:1170-1184.
1457 doi:10.1016/j.theriogenology.2005.02.003

Majeed KA, Rehman HU, Yousaf MS, Zaneb H, Rabbani I, Tahir SK and Rashid MA, 2017. Sub-chronic exposure to low concentration of dibutyl phthalate affects anthropometric parameters and markers of obesity in rats. Environmental Science and Pollution Research, 24:25462-25467. doi:10.1007/s11356-017-9952-y

Martin MM, Wu SM, Martin ALA, Rennert OM and Chan WY, 1998. Testicular seminoma in a patient with a constitutively activating mutation of the luteinizing hormone/chorionic gonadotropin receptor. European Journal of Endocrinology, 139:101-106. doi:10.1530/eje.0.1390101

Meeker JD, Calafat AM and Hauser R, 2009. Urinary Metabolites of Di(2-ethylhexyl) Phthalate Are Associated With Decreased Steroid Hormone Levels in Adult Men. Journal of Andrology, 30:287-297. doi:10.2164/jandrol.108.006403

Meeker JD and Ferguson KK, 2011. Relationship between Urinary Phthalate and Bisphenol A Concentrations and Serum Thyroid Measures in US Adults and Adolescents from the National Health and Nutrition Examination Survey (NHANES) 2007-2008. Environmental Health Perspectives, 119:1396-1402. doi:10.1289/ehp.1103582

Nakano T, Hurn PD, Herson PS and Traystman RJ, 2010. Testosterone exacerbates neuronal damage following cardiac arrest and cardiopulmonary resuscitation in mouse. Brain Res, 1357:124-130. doi:10.1016/j.brainres.2010.08.013

National Academies of Sciences, Engineering, and Medicine. 2017. Application of Systematic Review Methods in an Overall Strategy for Evaluating Low-Dose Toxicity from Endocrine Active Chemicals. Washington, DC: The National Academies Press. https://doi.org/10.17226/24758.

Oudir M, Chader H, Bouzid B, Bendisari K, Latreche B, Boudalia S and Iguer-Ouada M, 2018. Male rat exposure to low dose of di(2-ethylhexyl) phthalate during pre-pubertal, pubertal and post-pubertal periods: Impact on sperm count, gonad histology and testosterone secretion. Reprod Toxicol, 75:33-39. doi:10.1016/j.reprotox.2017.11.004

Pan G, Hanaoka T, Yu L, Na J, Yamano Y, Hara K, Ichiba M, Nakadate T, Kishi R, Wang P, Yin H, Zhang S and Feng Y, 2011. Associations between hazard indices of di-n-butylphthalate- and di-2-ethylhexylphthalate exposure and serum reproductive hormone levels among occupationally exposed and unexposed Chinese men. International Journal of Andrology, 34:E397-E406. doi: 10.1111/j.1365-2605.2011.01201.x

Philippat C, Mortamais M, Chevrier C, Petit C, Calafat AM, Ye XY, Silva MJ, Brambilla C, Pin I, Charles MA, Cordier S and Slama R, 2012. Exposure to Phthalates and Phenols during Pregnancy and Offspring Size at Birth. Environmental Health Perspectives, 120:464-470. doi:10.1289/ehp.1103634

Pocar P, Fiandanese N, Secchi C, Berrini A, Fischer B, Schmidt JS, Schaedlich K and Borromeo V, 2012. Exposure to Di(2-ethyl-hexyl) phthalate (DEHP) in Utero and during Lactation Causes Long-Term Pituitary-Gonadal Axis Disruption in Male and Female Mouse Offspring. Endocrinology, 153:937-948. doi:10.1210/en.2011-1450

Qi C, Ji X, Zhang G, Kang Y, Huang Y, Cui R, Li S, Cui H and Shi G, 2018. Haloperidol ameliorates androgen-induced behavioral deficits in developing male rats. J Endocrinol, 237:193-205. doi:10.1530/JOE-17-0642

1500 Repouskou A, Panagiotidou E, Panagopoulou L, Bisting PL, Tuck AR, Sjodin MOD, Lindberg J,
1501 Bozas E, Ruegg J, Gennings C, Bornehag CG, Damdimopoulou P, Stamatakis A and Kitraki
1502 E, 2019. Gestational exposure to an epidemiologically defined mixture of phthalates leads
1503 to gonadal dysfunction in mouse offspring of both sexes. Scientific Reports, 9.
1504 doi:10.1038/s41598-019-42377-6

1505 Sandy EH, Yao J, Zheng SX, Gogra AB, Chen HL, Zheng H, Yormah TBR, Zhang X, Zaray G,
1506 Ceccanti B and Choi MMF, 2010. A comparative cytotoxicity study of isomeric alkylphthalates
1507 to metabolically variant bacteria. Journal of Hazardous Materials, 182:631-639.
1508 doi:10.1016/j.jhazmat.2010.06.079

1509 Santangeli S, Maradonna F, Zanardini M, Notarstefano V, Gioacchini G, Forner-Piquer I, Habibi
1510 H and Carnevali O, 2017. Effects of diisononyl phthalate on Danio rerio reproduction.
1511 Environmental Pollution, 231:1051-1062. doi:10.1016/j.envpol.2017.08.060

1512 Sarachana T, Xu MY, Wu RC and Hu VW, 2011. Sex Hormones in Autism: Androgens and
1513 Estrogens Differentially and Reciprocally Regulate RORA, a Novel Candidate Gene for
1514 Autism. Plos One, 6. doi:10.1371/journal.pone.0017116

1515 Schwarz E, Guest PC, Rahmoune H, Wang L, Levin Y, Ingudomnukul E, Ruta L, Kent L, Spain
1516 M, Baron-Cohen S and Bahn S, 2011. Sex-specific serum biomarker patterns in adults with
1517 Asperger's syndrome. Mol Psychiatry, 16:1213-1220. doi:10.1038/mp.2010.102

1518 Stroustrup A, Bragg JB, Andra SS, Curtin PC, Spear EA, Sison DB, Just AC, Arora M and
1519 Gennings C, 2018. Neonatal intensive care unit phthalate exposure and preterm infant
1520 neurobehavioral performance. Plos One, 13. doi: 10.1371/journal.pone.0193835

1521 Takahashi K, Hosoya T, Onoe K, Takashima T, Tanaka M, Ishii A, Nakatomi Y, Tazawa S,
1522 Takahashi K, Doi H, Wada Y and Watanabe Y, 2018. Association between aromatase in
1523 human brains and personality traits. Scientific Reports, 8. doi:10.1038/s41598-018-35065-
1524 4

1525 Wang R, Xu XH, Weng HF, Yan SY and Sun YY, 2016a. Effects of early pubertal exposure to
1526 di-(2-ethylhexyl) phthalate on social behavior of mice. Hormones and Behavior, 80:117-
1527 124. doi:10.1016/j.yhbeh.2016.01.012

1528 Wang R, Xu XH and Zhu QJ, 2016b. Pubertal exposure to di-(2-ethylhexyl) phthalate influences
1529 social behavior and dopamine receptor D2 of adult female mice. Chemosphere, 144:1771-
1530 1779. doi:10.1016/j.chemosphere.2015.10.062

1531 Wang X, Wang L, Zhang JF, Yin WJ, Hou J, Zhang YJ, Hu C, Wang GY, Zhang R, Tao Y and
1532 Yuan J, 2018. Dose-response relationships between urinary phthalate metabolites and
1533 serum thyroid hormones among waste plastic recycling workers in China. Environmental
1534 Research, 165:63-70. doi:10.1016/j.envres.2018.04.004

1535 Zhao Y, Ao H, Chen L, Sottas CM, Ge RS, Li LX and Zhang YH, 2012. Mono-(2-ethylhexyl)
1536 phthalate affects the steroidogenesis in rat Leydig cells through provoking ROS
1537 perturbation. Toxicology in Vitro, 26:950-955. doi:10.1016/j.tiv.2012.04.003

1538 Zhao J, Ren S, Liu CY, Huo L, Liu Z and Zhai LL, 2018. Di-(2-Ethylhexyl) Phthalate Increases
1539 Obesity-Induced Damage to the Male Reproductive System in Mice. Oxidative Medicine and
1540 Cellular Longevity. doi: 10.1155/2018/1861984

Zhou M, Ford B, Lee D, Tindula G, Huen K, Tran V, Bradman A, Gunier R, Eskenazi B, Nomura DK and Holland N, 2018. Metabolomic Markers of Phthalate Exposure in Plasma and Urine of Pregnant Women. Frontiers in Public Health, 6. doi:10.3389/fpubh.2018.00298

Abbreviations

AOP	adverse outcome pathway
CEP	EFSA Panel on Food Contact Materials, Enzymes and Processing Aids
PPAR	peroxisome proliferator-activated receptor
CP	checkpoint
BBP	benzyl butyl phthalate
DBP	Dibutyl phthalate
DINP	diisononyl phthalate
DEHP	Bis(2-ethylhexyl) phthalate
ECHA	European Chemicals Agency
E2	estradiol
ED	endocrine disruptor
FCM	food contact materials
HBGV	health-based guidance values
HPG	hypothalamic-pituitary-gonadal axis
FSH	follicle-stimulating hormone
LH	luteinizing hormone
GD	gestational day
KER	key event relationship
LC	Leydig cells
MIE	molecular initiating event
MOA	mode of action
MTD	maximum tolerated dose
MOE	margin of exposure
NAMs	new approach methodologies
MEHP	monoethylhexyl phthalate
NMDR	non-monotonic dose-response
MDR	monotonic dose-response
NMDRC	non-monotonic dose-response curve
NOAEL	no observed adverse effect level
P_{NMDR}	probability of non-monotonic dose-response
PND	post-natal day
RA	risk assessment
RP	reference point

SC	Scientific Committee
SR	systematic review
SVHC	substances of very high concern
T	testosterone
ToR	terms of reference
TDI	tolerable daily intake
WG	working group