

ADOPTED: 22 September 2021

doi: 10.2903/j.efsa.2021.6877

Opinion on the impact of non-monotonic dose responses on EFSA's human health risk assessments

EFSA Scientific Committee,

Simon More, Diane Benford, Susanne Hougaard Bennekou, Vasileios Bampidis, Claude Bragard, Thorhallur Halldorsson, Antonio Hernandez-Jerez, Kostas Koutsoumanis, Claude Lambré, Kyriaki Machera, Ewen Mullins, Søren Saxmose Nielsen, Josef Schlatter, Dieter Schrenk, Dominique Turck, Jose Tarazona and Maged Younes

Abstract

This Opinion assesses the biological relevance of the non-monotonic dose responses (NMDR) identified in a previous EFSA External Report (Beausoleil et al., 2016) produced under GP/EFSA/SCER/2014/01 and the follow-up probabilistic assessment (Chevillotte et al., 2017a,b), focusing on the *in vivo* data sets fulfilling most of the checkpoints of the visual/statistical-based analysis identified in Beausoleil et al. (2016). The evaluation was completed with cases discussed in EFSA assessments and the update of the scientific literature. Observations of NMDR were confirmed in certain studies and are particularly relevant for receptor-mediated effects. Based on the results of the evaluation, the Opinion proposes an approach to be applied during the risk assessment process when apparent non-monotonicity is observed, also providing advice on specific elements to be considered to facilitate the assessment of NMDR in EFSA risk assessments. The proposed approach was applied to two case studies, Bisphenol A and bis(2-ethylhexyl phthalate (DEHP) and these evaluations are reported in dedicated annexes. Considering the potential impact of NMDRs in regulatory risk assessment, the Scientific Committee recommends a concerted international effort on developing internationally agreed guidance and harmonised frameworks for identifying and addressing NMDRs in the risk assessment process.

© 2021 European Food Safety Authority. *EFSA Journal* published by John Wiley and Sons Ltd on behalf of European Food Safety Authority.

Keywords: non-monotonic dose response (NMDR), statistical analysis, probabilistic analysis, biological relevance, reference dose, human health risk assessment

Requestor: EFSA Scientific Committee

Question number: EFSA-Q-2019-00530

Correspondence: sc.secretariat@efsa.europa.eu

Panel members: Simon More, Diane Benford, Susanne Hougaard Bennekou, Vasileios Bampidis, Claude Bragard, Thorhallur Halldorsson, Antonio Hernandez-Jerez, Kostas Koutsoumanis, Claude Lambré, Kyriaki Machera, Ewen Mullins, Hanspeter Naegeli (until July 2021), Søren Saxmose Nielsen, Josef Schlatter, Vittorio Silano (deceased), Dieter Schrenk, Dominique Turck and Maged Younes.

Declarations of interest: The declarations of interest of all scientific experts active in EFSA's work are available at <https://ess.efsa.europa.eu/doi/doiweb/doisearch>.

Acknowledgements: The Scientific Committee wishes to thank the following for the support provided to this scientific output: the hearing experts Laurent Bodin and Richard Sharp; the EFSA Staff Maria Chiara Astuto, Bernard Bottex, Anna Castoldi, Irene Cattaneo, Arianna Chiusolo, Andrea Terron, and Katharina Volk for the support provided to this scientific output.

Suggested citation: EFSA Scientific Committee, More S, Benford D, Hougaard Bennekou S, Bampidis V, Bragard C, Halldorsson T, Hernandez-Jerez A, Koutsoumanis K, Lambré C, Machera K, Mullins E, Nielsen SS, Schlatter J, Schrenk D, Turck D, Tarazona J and Younes M, 2021. Opinion on the impact of non-monotonic dose responses on EFSA's human health risk assessments. *EFSA Journal* 2021;19(10):6877, 68 pp. <https://doi.org/10.2903/j.efsa.2021.6877>

ISSN: 1831-4732

© 2021 European Food Safety Authority. *EFSA Journal* published by John Wiley and Sons Ltd on behalf of European Food Safety Authority.

This is an open access article under the terms of the [Creative Commons Attribution-NoDerivs](https://creativecommons.org/licenses/by/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited and no modifications or adaptations are made.

Reproduction of the images listed below is prohibited and permission must be sought directly from the copyright holder:

Figure 2: © Perharic et al., 2013.



The EFSA Journal is a publication of the European Food Safety Authority, a European agency funded by the European Union.



Table of contents

Abstract.....	1
1. Introduction.....	4
1.1. Background and Terms of Reference as provided by EFSA.....	6
1.2. Interpretation of the Terms of Reference.....	8
2. Data and methodologies.....	8
2.1. Data.....	8
2.2. Methodologies.....	9
3. Assessment.....	9
3.1. <i>In vivo</i> studies with data sets fulfilling five or six checkpoints.....	10
3.2. Other studies.....	17
3.2.1. Tropane alkaloids.....	17
3.2.2. Bisphenol A (BPA).....	18
3.2.3. Phthalates.....	19
3.3. Impact of NMDR on the risk assessment process.....	19
4. Conclusions.....	21
5. Recommendations.....	22
References.....	22
Abbreviations.....	26
Appendix A – Results from the additional probabilistic assessments.....	28
Annex A – Assessment of non-monotonicity reported for BPA.....	32
Annex B – Assessment of non-monotonicity reported for phthalates.....	44

1. Introduction

Since 2012, EFSA has conducted several activities for addressing the role of non-monotonic dose responses (NMDR) in the risk assessment process. These activities include a Scientific Colloquium (EFSA, 2012) as well as an outsourced procurement project to perform a systematic literature review covering the food and feed area and the optimisation of statistical approaches (Beausoleil et al., 2016). As a contribution to the international discussions on this topic, the Scientific Committee has analysed the impact of non-monotonicity in the human health risk assessment under EFSA remit, providing advice to the EFSA Panels when dealing with this situation during their assessments.

In a review on low-dose effects of endocrine active substances Vandenberg et al. (2012) provided the following definition of nonmonotonicity *'the slope of the dose-response curve changes sign from positive to negative or vice versa at some point along the range of doses examined'*. This definition also forms the basis for our evaluation; however, the definition does not cover how to formally assess, when relying on empirical data if there is a non-monotonic dose-response. This is a critical element when evaluating toxicity studies in the risk assessment process. The first step is to assess if there is any momentum in the data that is significantly different from the NULL hypothesis (this step is common to both monotonic and non-monotonic curves); and then, provided the data contain sufficient information, the shape of the dose response is evaluated through mathematical modelling.

Even if some form of non-monotonicity is detected statistically, determining if non-monotonicity is likely to be present is not always straightforward as often both non-monotonic and monotonic functions fit the data comparably well, and in other cases, one dose group may be driving the apparent non-monotonicity. In such cases, biological plausibility or absence of it is equally important to statistical considerations. Furthermore, when performing risk assessments, the presence of a statistically significant dose response is on its own not a sufficient condition when setting health-based guidance values as effect size and role in the adversity pathway are key attributes when evaluating risk.

As the title of this Opinion indicates, this document not only assesses the presence of any non-monotonicity statistically but also its biological plausibility and ways of addressing the presence of non-monotonic dose curves when performing risk assessment.

Using the definition of Vandenberg et al. (2012), the presence of NMDR implies that the slope of the dose-response curve changes sign at least once (see Figure 1). This may be explained, e.g. by a change in direction linked to toxicokinetics, toxicodynamics or both; as well as by superimposition of two or more underlying effects. In principle, non-monotonicity (i.e. the change of sign in the slope) may occur in different regions of the dose-response curve. Several other authors have reviewed the concept and consequences of non-monotonic dose responses (e.g. Beausoleil et al., 2013; Beausoleil et al., 2016; Hill et al., 2018; Zarn et al., 2020). Non-monotonicity occurring at the lower end of the dose-response has often been referred to as low-dose effects.¹ Several studies have reported non-monotonic dose-response curves for a number of chemicals, including pesticides, polychlorinated biphenyls (PCBs), dioxins and food contact materials such as bisphenol A (BPA) and phthalates, mainly regarding their endocrine activity (EFSA, 2012).

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) Scientific Colloquium.

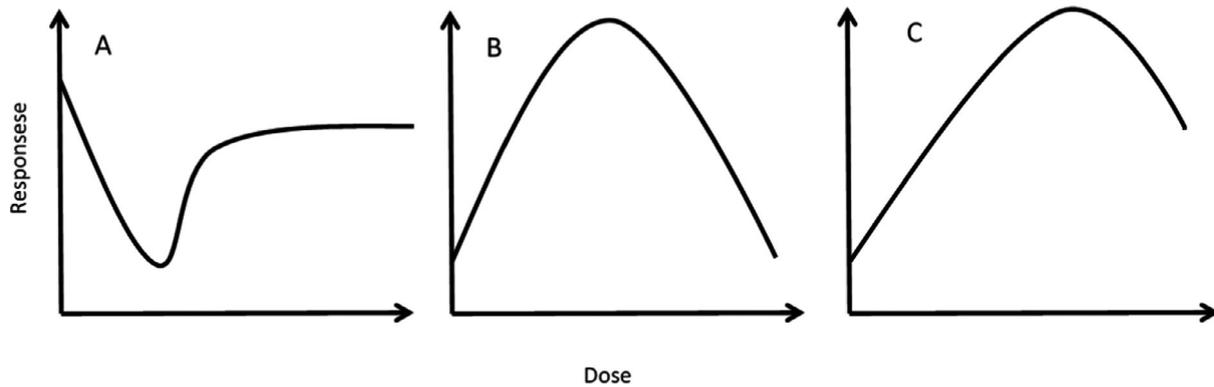


Figure 1: Examples of non-monotonic dose-response. The left figure (A) is one example of non-monotonicity occurring at the lower end of the dose-response. The middle figure (B) is an example of inverted U-shape dose-response, while the figure to the right (C) gives an example of non-monotonicity occurring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

Concepts describing NMDR have been debated 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 vs. 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). Conolly and Lutz (2004) described four examples of non-monotonic dose-response relations that they considered as superimposition of monotonic dose-responses of components of the respective biological system, and noted that numerous additional mechanisms can be proposed.

To discuss issues around low-dose effects and non-monotonic dose-response and their potential impact on risk assessment, EFSA organised 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 NMDRC² should not be disregarded in risk assessment, whereas others underscored the necessity to understand the mode of action before drawing conclusions for risk assessment. Thus, implementation of "low-dose effects" and NMDRCs in risk assessment strategies presents a scientific challenge and development of intelligent testing strategies to deal with these phenomena is necessary.'*

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). In that Report the scientific evidence for such NMDRs was also assessed. The systematic review, with two experts reviewing each data set, was performed in line with the EFSA guidance (EFSA, 2010).

Beausoleil et al. (2016) extracted dose-response data sets from studies having at least five 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, Beausoleil et al. (2016) proposed to use a set of six checkpoints as a tool for evaluating the evidence of NMDR in a single data set. 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 NMDRC: non-monotonic dose-response curve(s).

- 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 data set. 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 data sets indicated that 6% of the *in vivo* data sets fulfilled all six checkpoints and 20% fulfilled five checkpoints.

In total, 202 *in vivo* data sets (from 49 studies), 311 *in vitro* data sets (from 91 studies) and 9 epidemiological/human data sets (from two studies) were identified. 179 *in vivo* and 13 *in vitro* dose-response data sets were analysed.³ For 23 *in vivo* data sets, there were data limitations and these could, therefore, not be analysed. None of the data sets from epidemiological/human studies were analysed. In most of the *in vivo* data sets, it was concluded that the apparent NMDR was likely caused by a single outlying dose group. That is, in total only 10 of the 179 *in vivo* data sets fulfilled all visual/statistics-based checkpoints, while five checkpoints were fulfilled by 36 *in vivo* data sets (corresponding to 20%). Beausoleil et al. (2016) 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 characterise 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 characterisation procedure based on inter-dose statistical comparisons. The study focused on demonstrating the general applicability of 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. Curves were characterised 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 characterised 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 characterise 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 provides a strong methodological platform to assess statistically the presence of NMDR, but 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 Chevillotte et al. (2017b) and concluded that their method provides a probabilistic and objective characterisation 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 by outsourcing a procurement project to perform a systematic review of the literature referring to non-monotonic responses and a review of the scientific evidence for such NMDRs; the strength of the evidence was characterised using visual/statistics-based

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

checkpoints (Beausoleil et al., 2016). In this review, in total, 202 *in vivo* datasets (from 49 studies), 311 *in vitro* data sets (from 91 studies) and 9 epidemiological/human data sets (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%). Chevillotte et al. (2017a,b) reviewed the same data using a different but complementary probabilistic approach. 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 Beausoleil et al. (2016) (EFSA External 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 products⁴ 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 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 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 <https://efsa.onlinelibrary.wiley.com/doi/epdf/10.2903/j.efsa.2018.5311>

- 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 Beausoleil et al. (2016) (EFSA External Report) 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 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 for some effects. 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 Scientific Committee, 2021). In such cases, the NMDR is explained by two distinct but overlapping biological processes, which existing risk assessment paradigms can easily address. IPCS (2002) and EFSA Scientific Committee (2021) 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 characterised by a low-dose adaptive, stimulation or beneficial effect and a high-dose inhibitory or toxic effect (e.g. Calabrese and Baldwin, 2001). The effect at the lower end of the dose response relationship could, e.g. 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 *in vivo* studies presented in Beausoleil et al. (2016) (EFSA External Report) and the follow-up probabilistic assessment (Chevillotte et al., 2017a,b); covering substances of relevance in the food and feed area. All studies fulfilling 5 or 6 checkpoints in the Report have been included in the assessment, as well as the probabilistic assessments for these data sets.

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 following indexes: SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC.	TS = (monotonic OR nonmonotonic OR non-monotonic) AND TS = (toxic* AND dose)	The search was complemented with the analysis of the references and citations of the retrieved publications	<ul style="list-style-type: none"> • 225 articles retrieved • of those 19 <i>in vivo</i> experimental studies were selected for further evaluation after 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 Beausoleil et al. (2016), BPA and two phthalates, di(2-ethylhexyl) phthalate (DEHP) and di-n-butyl phthalate (DBP), were the substances under EFSA remit with the highest number of *in vivo* data sets reporting potential NMDR (35 for BPA, 30 for DEHP and 5 for DBP). However, the six checkpoints were met only for one of these data sets, aromatase activity in rats exposed to DEHP (Andrade et al., 2006). Considering the concordance between Beausoleil et al. (2016) and the complementary search, additional assessments regarding NMDR reported 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 Panel, 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 Beausoleil et al. (2016) and in the probabilistic assessment (Chevillotte et al., 2017a,b) has been briefly summarised 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; Chevillotte et al., 2017a,b), the results from the visual/statistical analysis of data sets judged to show potential NMDR (≥ 5 checkpoints) by Beausoleil et al. (2016) were compared with the probabilistic analysis conducted according to the methodology proposed by Chevillotte et al. (2017a,b). This probabilistic assessment methodology has been also applied to additional data sets selected from EFSA assessments and publications retrieved in the complementary literature search.

The biological relevance of potential NMDRs identified was assessed by expert judgement, analysing each selected publication. The systematic approach developed considered three key elements: (a) the extent to which the measured effect relates to the mode of action/Adverse Outcome Pathway, 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.⁵

This Scientific Opinion was published for public consultation, and the comments received have been assessed by the Working Group during the finalisation of this Scientific Opinion and published as a Technical Report (EFSA, 2021).

3. Assessment

The assessment is divided in two sections. Section 3.1 covers the *in vivo* studies included in Beausoleil et al. (2016) and containing data sets that fulfil five or six of the checkpoints in the visual/

⁵ Information on the endpoints selected as RP in EFSA assessments is available in the EFSA OpenFoodTox database. Available online: <https://www.efsa.europa.eu/en/data/chemical-hazards-data>

statistical analysis. Section 3.2 discusses other studies identified as potentially relevant from other EFSA activities but not covered in Beausoleil et al. (2016), and summarises the evaluations done for BPA and phthalates, which are detailed in Annexes A and B, respectively. One data set from Beausoleil et al. (2016) 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 data sets fulfilling five or six checkpoints

This section briefly describes examples of data sets from Beausoleil et al. (2016) 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 summarised in tables below. More specifically, Table 2 presents the cases in Beausoleil et al. (2016) meeting five or six checkpoints with well-defined biological explanation for NMDR, while Table 3 presents the assessment of the data sets 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 data set with possible NMDR is analysed regarding biological plausibility and role in adversity.

The examples from Beausoleil et al. (2016) cover 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.

Another group of studies summarised in Table 2 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 e.g. the review by Ferré et al. (2018) on the modes of action for the induction and inhibition of locomotor activity by caffeine). The capacity of nicotine to both activate and desensitise/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 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. The biological plausibility of NMDR in the area of developmental neurotoxicity (DNT) has previously been addressed in the NAFTA DNT Guidance (Moser et al., 2016). Biologically plausible observations are confirmed for assessment of motor activity and auditory startle reflex. 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 2: Studies fulfilling five or six 'checkpoints' in Beausoleil et al. (2016) (EFSA External Report) for which a well-defined biological explanation for NMDR could be identified

Publication, chemical, and measured effects	Dose range, No. of dose-groups (N) excluding controls	1) Presence/ shape of NMDR [□] (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
<p>Dey et al. (2009). Impact of resveratrol on indomethacin-induced gastric ulcer in mice</p> <p>1) Ulcer index 2) Myeloperoxidase (MPO) activity</p>	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 with 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)	<p>Ulcer index and MPO were measured at different time points, probability values are reported for one time point.</p> <p>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.</p>
<p>Takeda et al. (2002). Impact of rosmarinic acid on freezing behaviour of mice exposed to a conditioned fear stress (inescapable electric foot shocks)</p> <p>1) Duration of immobility</p>	0.25–4 mg/kg i.p. single dose N = 5	1) Yes U (CP-5)	1) Apical effect	1) Yes	1) Unclear, is an alteration of the natural response to stress. Spontaneous motor activity was not affected.	1) P _{NMDR} 78.35 (result after 3 days)	<p>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.</p>

Publication, chemical, and measured effects	Dose range, No. of dose-groups (N) excluding controls	1) Presence/ shape of NMDR [□] (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
Buenafe et al. (2013). Anticonvulsant activity of Tanshinone IIA in mice subjected to electrical stimulus through the corneas. 1) Number of mice protected	0.1–10 mg/kg i.v. N = 5	1) Yes ∩ (CP-5)	1) Apical effect	1) Yes	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 i.v. Biphasic/hormetic dose responses have indeed been previously reported in chemically diverse pro- and anticonvulsant agents with different modes of action.
Haldner et al. (2004). Impact of caffeine on locomotor activity in mice 1) Horizontal activity (number of counts indicating movements to adjacent cells)	3.75–100 mg/kg i.p. 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 i.p. 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 mobilisation of calcium from intracellular stores (Fisone et al., 2004)

Publication, chemical, and measured effects	Dose range, No. of dose-groups (N) excluding controls	1) Presence/ shape of NMDR [¶] (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
Zhang et al. (2011). Impact of caffeine on locomotor activity in mice 1) Horizontal activity (travel distance) 2) Distance ratios in central and peripheral regions	1–100 mg/kg i.p. 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.
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 i.p. 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.

Publication, chemical, and measured effects	Dose range, No. of dose-groups (N) excluding controls	1) Presence/ shape of NMDR [¶] (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
Varlinskaya and Spear (2009). Impact of ethanol on motor activity in mice 1) Behaviour as social investigation 2) Behaviour as play fighting	0.25–1.25 g/kg s.c. 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 on play fighting but not on social investigation.
Bai and Zhu (2010). role of two bioflavonoids as co-substrates for cyclooxygenases (COX) in rats 1) Impact of myricetin on PGE2 levels plasma 2) Impact of quercetin 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 Beausoleil et al. (2016):

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 an NMDR with U (or J) shape, the symbol ∩ indicates an NMDR with inverted U (or J) shape.

‡: Only addressed when a possible NMDR is confirmed under 1. Presence/shape of NMDR.

‡: The key Monte Carlo resampling results are presented, for additional results see the publications.

Table 3: Other studies fulfilling five or six 'checkpoints' in Beausoleil et al. (2016) (EFSA External Report) for which well-defined biological explanation for NMDR were subject to some uncertainty

Publication, chemical and measured effect	Dose range, No. 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
Puatanachokchai 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) 2 α -testosterone hydroxylase activity in liver 5) 8OHdG formation in liver 6) NADPH-P450 reductase activity in liver 7) 16 α -testosterone hydroxylase activity in liver	0.01–500 mg/kg diet 10 week (0.001–50 mg/kg bw) N = 7 All rats had received 100 mg/kg bw i.p. diethylnitrosamine weekly 3 times before starting alpha HCH exposure	1) No (CP-3) 2) Yes U (CP-3) 3) Yes U (none) 4) Yes \cap (none) 5) Yes U (none) 6) Yes U (none) 7) Yes \cap (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 (\cap)	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 i.p. on rats 1) Protein expression in cerebral cortex as marker for stress response	2–10 mg/kg bw i.p., 1x N = 6	1) Yes, \cap but toxicity could explain the decrease in protein expression at doses > 6 mg/kg (none)	1) Early event	1) Yes	1) Unclear	1) P _{NMDR} 72.0	Not relevant for the much lower human exposure. Furthermore, acute i.p. application
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 4 week 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	

Publication, chemical and measured effect	Dose range, No. 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
Sukata et al. (2002). Effects of DDT on rats. 1) Proliferation of GST-P positive hepatic foci 2) mRNA IL-1 receptor type 1 (Figure 3)	0.5–500 mg/kg diet 16 week (0.05–20 mg/kg bw) N = 8	1) No (CP-3) 2) Yes, trend, not stat. Sign. \cap (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 \cap
Wildemann et al. (2015) Effects of lead acetate on rats 1) Body weight gain 2) Pulse pressure	0.004–45 mg/kg bw per day N = 8 Drinking water	1) Yes, \cap (CP-5) 2) Yes, U	1) Apical 2) Intermediate	?	1) Yes, body weight gain was 113 g control vs. up to 224 g treated 2–7 Yes	1) P_{NMDR} 92.38 2) P_{NMDR} 76.64	All the haemodynamic effects are linked. Other possible non-monotonic responses but with less than 5 checkpoints observed for Systolic blood pressure Stroke volume Cardiac output
Zorrilla et al. (2009) Effects of triclosan on juvenile rats 1) Triiodothyronine (T3) serum	3–300 mg/kg per day N = 5 Gavage	1) Yes, \cap (CP-3)	1) Intermediate	1) ?	1) Yes, reduction in T3 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. The main effect is for T4 and is clearly monotonic.

*: CP = checkpoint as defined in Beausoleil et al. (2016):

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 an NMDR with U (or J) shape, the symbol \cap indicates an NMDR with inverted U (or J) shape.

‡: Only addressed when a possible NMDR is confirmed under 1. Presence/shape of NMDR.

†: The key Monte Carlo resampling results are presented, for additional results, see the publications.

Consistency between the two methodologies for the statistical assessment of NMDR (Beausoleil et al., 2016; Chevillotte et al., 2017a,b) 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 Chevillotte et al. (2017a,b) was higher than 78% in all cases, and the NMDR confirmed by the expert judgement.

For the studies reviewed (Tables 2 and 3) 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 met for some data sets 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. As summarised in Table 3, in two cases (proliferation of GST-P positive hepatic foci in Puatanochochai et al. (2016), and in Sukata et al. (2002), the expert judgement concluded that there were no indications for NMDR, despite the data set fulfilled five checkpoints and the likelihood in the probabilistic analysis was higher than 75%. The biological plausibility was clear for all data sets reported in Table 2, but remained doubtful for the majority of data sets reported in Table 3.

3.2. Other studies

3.2.1. Tropane alkaloids

Tropane alkaloids were identified from an Opinion of the EFSA Scientific Panel on Contaminants in the Food Chain (EFSA CONTAM Panel, 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 decrease in 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.

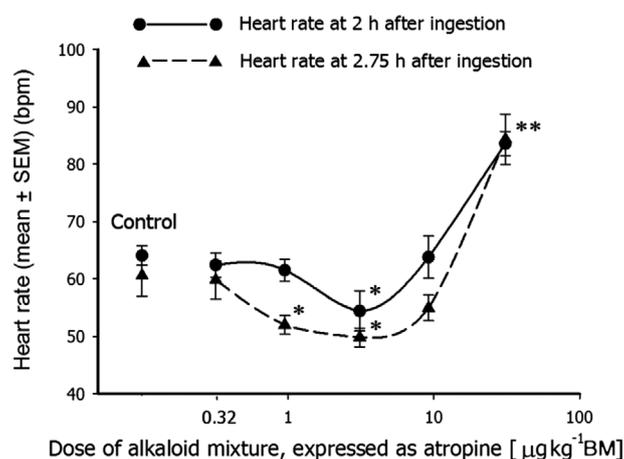


Figure 2: Dose-response curve for heart rate vs. the dose of the atropine/scopolamine mixture, expressed as atropine (* $p < 0.005$, ** $p < 0.001$). Reproduced with permission from Perharič et al. (2013) (DOI 10.1002/jat.2797)

Results of additional probabilistic assessments for Perharič et al. (2013) conducted according to the methodology proposed by Chevillotte et al. (2017a,b) confirm the NMDR with associated probabilities for a U-shaped dose-response of 66.1% and 86.7% at 2 and 2.75 h, respectively, for Monte Carlo resampling and almost 100% for Latin-Hypercube resampling (see Appendix A for the detailed results).

3.2.2. Bisphenol A (BPA)

Beausoleil et al. (2016) identified four studies on BPA where possible NMDR had been reported. These studies were not necessarily picked up because they provided convincing evidence of NMDR but rather because the word came up in the publication. As an example, a study by Tyl et al. (2002), was identified as the study was designed to examine possible NMDR for developmental effects of BPA. Although the authors concluded in their publication that no indication of NMDR was present in their results, the study was retrieved as the word 'NMDR' was identified in their literature search and liver weight in the second generation (F2) was still evaluated by Beausoleil et al. (2016). The results being in line with those of the authors that the presence of NMDR was unclear (only three checkpoints were fulfilled). The SC evaluation reached the same conclusion (See Annex A). Other studies on BPA identified by Beausoleil et al. (2016) included possible NMDR for extracellular kinase signalling in cerebellar cortex (pERK-IRCellAtP10) (Zsarnovszky et al., 2005), semen quality (Kendig et al., 2012) and gonadal and renal fat pads (Angle et al., 2013). Only four checkpoints were fulfilled for each of these studies. For risk assessment, the relevance of an effect on extracellular kinase signalling in 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 indicating 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 and few additional studies were included based on suggestions received from the public consultation process. 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). Another study reported a possible U-shaped NMDR for sperm count (Hass et al., 2016) with modest effect size, but the probabilistic assessment concluded that the dose response was more likely to be monotonic with high probability (97%, See Annex A). Based on findings reported in these studies, 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. A probabilistic assessment conducted according to Chevillotte et al. (2017a,b) could not confirm that conclusion (probability for a NMDR 58.8% for Monte Carlo resampling, while 45.9% for Latin Hypercube resampling (see Appendix A for details). Furthermore, no signs of NMDR or any differences in body weight were observed at later ages, suggesting that this finding could be an outlier. Overall, the possible NMDR reported in the above-mentioned studies on measures of body composition seems weak as seen by high variability across dose groups and modest effect size.

Using a transgenic mouse model (MMTV-erbB2) with a high rate of spontaneous tumorigenesis, Jenkins et al. (2011) reported an NMDR between low dose (2.5, 25, 250 and 2,500 µg/kg bw) BPA exposure and mammary cancer. In general, experimental studies in transgenic animals are used for mechanistic insights but as such are not used on their own to identify a reference point to be used in setting a HBGV or MOE. Therefore, this study was not included in our assessment. Furthermore, the SC noted some deficiencies such as the number of animals per dose group in this study was highly uneven (n between 37 and 94).

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 (Montévil et al., 2020), modest changes in % basophil and serum bile acid concentrations (Badding et al., 2019). A detailed assessment of these studies is provided in Annex A. 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.

3.2.3. Phthalates

Beausoleil et al. (2016) (EFSA External Report) identified, using the checkpoint approach, an NMDR for DEHP on aromatase activity, and there are a number of publications suggesting NMDR for phthalates and DEHP in particular. The assessment of those studies, included in Annex B, 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 Beausoleil et al. (2016) for aromatase as this enzyme is involved in testosterone metabolism.

While a monotonic decrease in fetal testosterone levels is observed following DEHP exposure, NMDR has been observed for postnatal testosterone measurements. The NMDR observed for this intermediate effect in postnatal situations could be related to several mechanisms that may run in parallel, including disturbances in steroidogenesis, or in the hypothalamic–pituitary–gonadal axis (HPG) feedback mechanism. The assessment of the available evidence includes biphasic responses following phthalate exposure in several steps linked to testosterone synthesis and metabolism, including an NMDR for aromatase activity; in addition, an NMDR for testosterone levels could result from monotonic disturbances of different steps in the steroidogenesis pathway. Another possible mechanistic interpretation could be the overstimulation of the feedback mechanisms following chronic exposures to low doses.

According to the proposed approach, as testosterone levels are an intermediate event, the next steps should be to assess the possible biological relevance of these effects, in particular if a (quantitative) relation between these effects and an adverse outcome (i.e. apical effect) can be established, ideally through a mechanistic sequence (AOP).

There is information indicating that postnatal increases in testosterone levels under certain conditions may trigger pathways resulting in adverse outcomes, although the relevance of these findings for experimental postnatal studies with phthalates is unclear. Examples cover 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) and apical effects associated with overexpression of androgens (Hotchkiss et al., 2007; Martin et al., 1998). There are also some epidemiological studies linking phthalate exposure with metabolomic alterations (Zhou et al., 2018, neurodevelopmental Braun, 2017, Engel et al., 2018) and effects attributable to hyperandrogenism (Colon et al., 2000). As mentioned, a comprehensive in depth assessment of these effects has not been performed, and is outside the scope of this Opinion.

3.3. Impact of 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 (i.e. a change of the sign of the slope) may occur in 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 if an NO(A)EL or LO(A)EL is used as the RP to establish a HBGV or to calculate an MOE. That is, the RP would not change because of more adverse effects occurring at higher doses. The SC acknowledges that the benchmark-dose (BMD) methodology in its current form⁶ should be used with caution in case of NMDR. Furthermore, if non-monotonicity occurs at high end of the dose–response curve, these effects are often caused by saturation 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. The identified NMDR

⁶ According to current EFSA guidance, the BMD approach is restricted to monotonic functions meaning that if effects showing NMDR would be used to derive a reference point the NOAEL/LOAEL approach would be used. Future update on the BMD guidances may include extension of the methodology to cover NMDR.

for non-nutrients observed *in vivo*⁷ may concern early or intermediate events in the toxicity pathways, with no indications of non-monotonicity in the related apical endpoints usually used for identifying an RP. During the evaluation of these NMDR, it is 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). Receptor-mediated effects provide additional examples: It is well established that compounds interacting with cellular receptors may lead to biphasic 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.

With the current design for guideline *in vivo* studies, the low number of doses represents a limitation for the identification and analysis of NMDR. To facilitate the assessment, and also minimise the need for repeating animal studies, New Approach Methodologies (NAM)-based studies should be considered. The integration of available animal and human studies with NAMs may provide the mechanistic understanding required for implementing the use of AOP approaches. Complementing the OECD AOP programme,⁸ several regulatory agencies including EFSA are investing in the use of mechanistic data for the development and validation of AOPs in order to inform the risk assessment process; several proof of concept case studies in the EFSA remit are available (e.g. Bal-Price et al., 2018) or ongoing. Quantitative AOPs are particularly promising for regulatory uses (e.g. Spinu et al., 2020), while AOP networks cover the need for addressing complex relationships (Knapen et al., 2018; Villeneuve et al., 2018). *In vitro* studies are more able to include a large number of doses (concentrations), covering a broad range of the dose-response curve and facilitating the identification of NMDRs. This integration (NAMs and available animal and human data) will facilitate the identification of NMDRs in early and intermediate events. Understanding the reason behind the NMDR (e.g. biphasic response of a receptor-mediated endpoint; combination of several modes of action, or toxicity at high doses) facilitates the assessment of its relevance for the risk assessment.

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 an 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

⁷ 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.

⁸ <https://www.oecd.org/chemicalsafety/testing/adverse-outcome-pathways-molecular-screening-and-toxicogenomics.htm>

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. As detailed above, NAMs would reduce the need for further animal studies and are 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 may be difficult to address. Inclusion of a sufficient number of dose groups would also benefit the application of the BMD approach. Furthermore, mechanistic data would inform whether or not early/intermediate effects show non-monotonicity.

4. Conclusions

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

- 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.
- In assessing dose-response relationships for non-monotonicity, the checkpoint approach may in some cases yield different result than those obtained through probabilistic (statistical) methodology;
- 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 a molecular initiating event or an early/intermediate event, the potential for propagating towards an apical effect needs to be investigated and checked for its biological relevance as above. It should be noted that molecular initiating events 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 Section 3.3. This approach is recommended for application in cases of apparent non-monotonicity.

Observations of NMDR have been confirmed in certain studies and are particularly relevant for receptor-mediated effects. In order to facilitate the assessment of NMDR, the Scientific Committee advises that the following points should be considered in EFSA risk assessments:

- 1) Evidence for non-monotonicity of apical effects should be assessed in terms of statistical rigor and biological plausibility. Indications of possible NMDR should be investigated and considered during the risk assessment process according to the process detailed in this Opinion.
- 2) The benchmark-dose (BMD) methodology in its current form should be used with caution when establishing RPs in the case of NMDRs.

- 3) It is recommended to explore the mechanistic basis of these NMDR (i.e. using NAM-based data) and to integrate the results in AOP-like approaches during the risk assessment process when needed.

The approach proposed in this Opinion was applied to two case studies: Bisphenol A (BPA) and Phthalates. Based on that work, no clear indications of NMDR were detected for BPA, while for the phthalate DEHP, indications for a biologically plausible NMDR were observed for an intermediate effect, postnatal testosterone levels, with several mechanisms supporting biological plausibility. The impact of this NMDR on the risk assessment of DEHP should be evaluated further.

5. Recommendations

- There is a need for an international effort to provide more detailed dose-response information for risk assessment, taking into account animal welfare considerations as well as developments in the field of NAMs. This would facilitate capturing and concluding on the presence of an NMDR.
- Considering the potential impact of NMDRs in regulatory risk assessment, the SC encourages a concerted international effort on developing:
 - a) internationally agreed guidance on the statistical approaches for identifying NMDR, and
 - b) harmonised frameworks for addressing NMDR in the risk assessment process.

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. <https://doi.org/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. *Reproductive Toxicology*, 42, 256–268. <https://doi.org/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 and Chemical Toxicology*, 131. <https://doi.org/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. <https://doi.org/10.1016/j.plefa.2009.10.006>
- Bal-Price A, Leist M, Schildknecht S, Tschudi-Monnet F, Paini A and Terron A, 2018. Adverse Outcome Pathway on Inhibition of the mitochondrial complex I of nigro-striatal neurons leading to parkinsonian motor deficits. <https://doi.org/10.1787/b46c3c00-en>
- 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, Ormsby JN, Gies A, Hass U, Heindel JJ, Holmer ML, Nielsen PJ, Munn S and Schoenfelder G, 2013. Low dose effects and non-monotonic dose responses for endocrine active chemicals: science to practice workshop: workshop summary. *Chemosphere*, 93, 847–856. <https://doi.org/10.1016/j.chemosphere.2013.06.043>
- 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. <https://doi.org/10.2903/sp.efsa.2016.EN-1027>
- Braun JM, 2017. Early-life exposure to EDCs: role in childhood obesity and neurodevelopment. *Nat Rev Endocrinol*, 13, 161–173. <https://doi.org/10.1038/nrendo.2016.186>
- 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. <https://doi.org/10.1021/cn400140e>
- Calabrese EJ, 2005. Toxicological awakenings: the rebirth of hormesis as a central pillar of toxicology. *Toxicology and Applied Pharmacology*, 204, 1–8. <https://doi.org/10.1016/j.taap.2004.11.015>
- Calabrese EJ and Baldwin LA, 2001. Hormesis: a generalizable and unifying hypothesis. *Critical Reviews in Toxicology*, 31, 353–424. <https://doi.org/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. <https://doi.org/10.1038/s41514-017-0013-z>

- 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. <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 and Chemical Toxicology*, 106, 376–385. <https://doi.org/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 and Chemical Toxicology*, 110, 214–228. <https://doi.org/10.1016/j.fct.2017.10.030>
- 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. <https://doi.org/10.2307/3434999>
- Conolly RB and Lutz WK, 2004. Nonmonotonic dose-response relationships: mechanistic basis, kinetic modeling, and implications for risk assessment. *Toxicological Sciences*, 77, 151–157. <https://doi.org/10.1093/toxsci/kfh007>
- Correa M, Arizzi MN, Betz A, Mingote S and Salamone JD, 2003. Open field locomotor effects in rats after intraventricular injections of ethanol and the ethanol metabolites acetaldehyde and acetate. *Brain Research Bulletin*, 62, 197–202. <https://doi.org/10.1016/j.brainresbull.2003.09.013>
- 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. *Toxicology and Applied Pharmacology*, 347, 1–9. <https://doi.org/10.1016/j.taap.2018.03.021>
- Dey A, Guha P, Chattopadhyay S and Bandyopadhyay SK, 2009. Biphasic activity of resveratrol on indomethacin-induced gastric ulcers. *Biochemical and Biophysical Research Communications*, 381, 90–95. <https://doi.org/10.1016/j.bbrc.2009.02.027>
- ECHA and EFSA (European Chemicals Agency and European Food Safety Authority) with the technical support of the Joint Research Centre (JRC), Andersson N, Arena M, Auteri D, Barmaz S, Grignard E, Kienzler A, Lepper P, Lostia AM, Munn S, Parra Morte JM, Pellizzato F, Tarazona J, Terron A and Van der Linden S, 2018. Guidance for the identification of endocrine disruptors in the context of Regulations (EU) No 528/2012 and (EC) No 1107/2009. *EFSA Journal* 2018;16(6):5311, 135 pp. <https://doi.org/10.2903/j.efsa.2018.5311>. ECHA-18-G-01-EN.
- EFSA (European Food Safety Authority), 2010. Application of systematic review methodology to food and feed safety assessments to support decision making. *EFSA Journal* 2010;8(6):1637, 90 pp. <https://doi.org/10.2903/j.efsa.2010.1637>
- EFSA (European Food Safety Authority), 2012. EFSA's 17th Scientific Colloquium on low dose response in toxicology and risk assessment. *EFSA Supporting Publication* 2012;9(11):EN-353, 64 pp. <https://doi.org/10.2903/sp.efsa.2012.EN-353>
- EFSA (European Food Safety Authority), 2021. Outcome of the public consultation on draft EFSA Opinion on the impact of non-monotonic dose responses on EFSA's human health risk assessments. *EFSA supporting publication* 2021;EN-6878, 109 pp. <https://doi.org/10.2903/sp.efsa.2021.EN-6878>
- 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, Rivière G, Steffensen I-L, Tlustos C, Van Loveren H, Vernis L, Zorn H, Cravedi J-P, Fortes C, Tavares Poças MF, Waalkens-Berendsen I, Wölfle 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>
- EFSA CONTAM Panel (EFSA Panel on Contaminants in the Food Chain), 2013. Scientific Opinion on Tropic alkaloids in food and feed. *EFSA Journal* 2013;11(10):3386, 113 pp. <https://doi.org/10.2903/j.efsa.2013.3386>
- EFSA Scientific Committee, More S, Bampidis V, Benford D, Bragard C, Halldorsson T, Hougaard Bennekou S, Koutsoumanis K, Machera K, Naegeli H, Nielsen S, Schlatter J, Schrenk D, Silano V, Turck D, Younes M, Aggett P, Castenmiller J, Giarola A, de Sesmaisons-Lecarré A, Tarazona J, Verhagen H and Hernández-Jerez A, 2021. Statement on the derivation of Health-Based Guidance Values (HBGVs) for regulated products that are also nutrients. *EFSA Journal* 2021;19(3):6479, 39 pp. <https://doi.org/10.2903/j.efsa.2021.6479>
- 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. <https://doi.org/10.1289/ehp2358>
- Escarabajal MD and Aragon CMG, 2002. The effect of cyanamide and 4-methylpyrazole on the ethanol-induced locomotor activity in mice. *Pharmacology, Biochemistry, and Behavior*, 72, 389–395. [https://doi.org/10.1016/s0091-3057\(01\)00762-6](https://doi.org/10.1016/s0091-3057(01)00762-6)

- Ferré S, Díaz-Ríos M, Salamone JD and Prediger RD, 2018. New developments on the adenosine mechanisms of the central effects of caffeine and their implications for neuropsychiatric disorders. *Journal of Caffeine and Adenosine Research*, 8, 121–130. <https://doi.org/10.1089/caff.2018.0017>
- Fisone G, Borgkvist A and Usiello A, 2004. Caffeine as a psychomotor stimulant: mechanism of action. *Cellular and Molecular Life Sciences*, 61, 857–872. <https://doi.org/10.1007/s00018-003-3269-3>
- Halldner L, Adén U, Dahlberg V, Johansson B, Ledent C and Fredholm BB, 2004. The adenosine A1 receptor contributes to the stimulatory, but not the inhibitory effect of caffeine on locomotion: a study in mice lacking adenosine A1 and/or A2A receptors. *Neuropharmacology*, 46, 1008–1017. <https://doi.org/10.1016/j.neuropharm.2004.01.014>
- 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. <https://doi.org/10.1111/andr.12176>
- Hill CE, Myers JP and Vandenberg LN, 2018. Nonmonotonic dose-response curves occur in dose ranges that are relevant to regulatory decision-making. *Dose-Response*, 16. <https://doi.org/10.1177/1559325818798282>
- 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 Acad Sciences, New York. pp. 272–282.
- Hotchkiss AK, Lambright CS, Ostby JS, Parks-Saldutti L, Vandenberg 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. <https://doi.org/10.1093/toxsci/kfm002>
- IPCS (International Programme on Chemical Safety), 2002. Principles and Methods for the Assessment of Risk from Essential Trace Elements. World Health Organization. Available online: <https://apps.who.int/iris/handle/10665/42416>
- IPCS (International Programme on Chemical Safety), 2009. Environmental Health Criteria 240, Principles and Methods for the Risk Assessment of Chemicals in Food. World Health Organization.
- Jenkins S, Wang J, Eltoum I, Desmond R and Lamartiniere CA, 2011. Chronic oral exposure to bisphenol A results in a nonmonotonic dose response in mammary carcinogenesis and metastasis in MMTV-erbB2 mice. *Environmental Health Perspectives*, 119, 1604–1609. <https://doi.org/10.1289/ehp.1103850>
- 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 17 α -ethinyl estradiol in CD1 mice. *Int J Toxicol*, 31, 537–550. <https://doi.org/10.1177/1091581812463254>
- Knapen D, Angrish MM, Fortin MC, Katsiadaki I, Leonard M, Margiotta-Casaluci L, Munn S, O'Brien JM, Pollesch N, Smith LC, Zhang X and Villeneuve DL, 2018. Adverse outcome pathway networks I: development and applications. *Environmental Toxicology and Chemistry*, 37, 1723–1733. <https://doi.org/10.1002/etc.4125>
- 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. <https://doi.org/10.1186/1476-069x-14-13>
- Marin MT, Zancheta R, Paro AH, Possi AP, Cruz FC and Planeta CS, 2011. Comparison of caffeine-induced locomotor activity between adolescent and adult rats. *European Journal of Pharmacology*, 660, 363–367. <https://doi.org/10.1016/j.ejphar.2011.03.052>
- 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. <https://doi.org/10.1530/eje.0.1390101>
- 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. <https://doi.org/10.1289/EHP6301>
- Moser V, Bailey F, Bowers W, Raffaele K, Crofton K and Gilbert M, 2016. NAFTA (North American Free Trade Agreement) Technical Working Group on Pesticides (TWG), Developmental Neurotoxicity Study (DNT) Guidance Document.
- Nakano T, Hurn PD, Herson PS and Traystman RJ, 2010. Testosterone exacerbates neuronal damage following cardiac arrest and cardiopulmonary resuscitation in mouse. *Brain Research*, 1357, 124–130. <https://doi.org/10.1016/j.brainres.2010.08.013>
- Perharič L, Juvan KA and Stanovnik L, 2013. Acute effects of a low-dose atropine/scopolamine mixture as a food contaminant in human volunteers. *Journal of Applied Toxicology*, 33, 980–990.
- Picciotto MR, Addy NA, Mineur YS and Brunzell DH, 2008. It is not “either/or”: activation and desensitization of nicotinic acetylcholine receptors both contribute to behaviors related to nicotine addiction and mood. *Progress in Neurobiology*, 84, 329–342. <https://doi.org/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. <https://doi.org/10.1007/BF00174872>

- Pitschner HF, Schulte B, Neuzner J, Wellstein A, Palm D and Schlepfer M, 1994. Subtypes of muscarinic receptors— aspects of their physiologic significance for controlling heart rate in the human. *Zeitschrift für Kardiologie*, 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. <https://doi.org/10.1016/j.canlet.2005.09.006>
- 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. *Journal of Endocrinology*, 237, 193–205. <https://doi.org/10.1530/JOE-17-0642>
- Schwarz E, Guest PC, Rahmoune H, Wang L, Levin Y, Ingudomnukul E, Ruta L, Kent L, Spain M, Baron-Cohen S and Bahn S, 2011. Sex-specific serum biomarker patterns in adults with Asperger's syndrome. *Molecular Psychiatry*, 16, 1213–1220. <https://doi.org/10.1038/mp.2010.102>
- 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. *Journal of Toxicological Sciences*, 34, 469–482. <https://doi.org/10.2131/jts.34.469>
- Spinu N, Cronin MTD, Enoch SJ, Madden JC and Worth AP, 2020. Quantitative adverse outcome pathway (qAOP) models for toxicity prediction. *Archives of Toxicology*, 94, 1497–1510. <https://doi.org/10.1007/s00204-020-02774-7>
- 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. <https://doi.org/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 (Berl)*, 164, 233–235. <https://doi.org/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. <https://doi.org/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. *Toxicological Sciences*, 68, 121–146. <https://doi.org/10.1093/toxsci/68.1.121>
- Uchtmann KS, Taylor JA, Timms BG, Stahlhut RW, Rieke EA, Ellersieck MR, vom Saal FS and Rieke 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. <https://doi.org/10.1016/j.reprotox.2019.11.007>
- Vandenberg LN, Colborn T, Hayes TB, Heindel JJ, Jacobs Jr DR, Lee D-H, Shioda T, Soto AM, vom Saal FS, Welshons WV, Zoeller RT and Myers JP, 2012. Hormones and endocrine-disrupting chemicals: low-dose effects and nonmonotonic dose responses. *Endocrine Reviews*, 33, 378–455. <https://doi.org/10.1210/er.2011-1050>
- 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. <https://doi.org/10.1111/j.1530-0277.2009.00920.x>
- Villeneuve DL, Angrish MM, Fortin MC, Katsiadaki I, Leonard M, Margiotta-Casaluci L, Munn S, O'Brien JM, Pollesch NL, Smith LC, Zhang X and Knapen D, 2018. Adverse outcome pathway networks II: network analytics. *Environmental Toxicology and Chemistry*, 37, 1734–1748. <https://doi.org/10.1002/etc.4124>
- 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. <https://doi.org/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. <https://doi.org/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. <https://doi.org/10.2203/dose-response.11-030.Yuanqing>
- Zarn JA, Zürcher UA and Geiser HC, 2020. Toxic responses induced at high doses may affect benchmark doses. *Dose-Response*, 18, 1559325820919605. <https://doi.org/10.1177/1559325820919605>
- 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. <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. <https://doi.org/10.2203/dose-response.11-055.Rongzhu>
- 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. <https://doi.org/10.3389/fpubh.2018.00298>
- 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. <https://doi.org/10.1016/j.envpol.2017.05.043>
- 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. <https://doi.org/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. <https://doi.org/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
i.p.	intraperitoneal administration
i.v.	intravenous administration
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
p.o.	oral administration
RA	risk assessment
RP	reference point
SC	Scientific Committee
s.c.	Subcutaneous administration

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

Appendix A – Results from the additional probabilistic assessments

Tables A.1 and A.2 show the results of additional probabilistic assessments for Perharič et al. (2013) conducted according to the methodology proposed by Chevillotte et al. (2017a,b).

Table A.1: From Perharič et al. (2013) (Table 4). Endpoint: Heart rate at 2 h

	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 A.2: From Perharič et al. (2013) (Table 4). Endpoint: Heart rate at 2.75 h

	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

Tables A.3, A.4 and A.5 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 A.3: 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 A.4: 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

	Probabilistic methodology Monte Carlo resampling	Probabilistic methodology Latin-Hypercube resampling
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

Table A.5: From Uchtmann et al. (2020) (Table 3 – Supplementary material). Endpoint urogenital 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

Tables A.6 and A.7 show the results of additional probabilistic assessments for Hass et al. (2016) conducted according to the methodology proposed by Chevillotte et al. (2017a,b).

Table A.6: From Hass et al., 2016 (Figure 2). Endpoint: sperm count in male offspring

	Probabilistic methodology Monte Carlo resampling	Probabilistic methodology Latin-Hypercube resampling
Type of dose-response	Prob (%)	Prob (%)
No DR	0.28	0
MDR increasing	48.06	100
MDR decreasing	0.03	0
NMDR U	48.18	0
NMDR inverted-U	0.01	0
NMDR complex	3.44	0
Total	100	100

Table A.7: From Hass et al. (2016) (Figure 4a). Endpoint: swim length in female offspring

	Probabilistic methodology Monte Carlo resampling	Probabilistic methodology Latin-Hypercube resampling
Type of dose-response	Prob (%)	Prob (%)
No DR	9.65	0
MDR increasing	10.36	0
MDR decreasing	18.26	0.02
NMDR U	58.28	99.98
NMDR inverted-U	0.19	0
NMDR complex	3.26	0
Total	100	100

Tables A.8, A.9, A.10, A.11, A.12, A.13 show the results of additional probabilistic assessments for Rubin et al. (2017) conducted according to the methodology proposed by Chevillotte et al. (2017a,b).

Table A.8: From Rubin et al. (2017) (Figure 1b). Endpoint body weight in female exposed perinatally and peripubertally, PND28

	Probabilistic methodology Monte Carlo resampling	Probabilistic methodology Latin-Hypercube resampling
Type of dose-response	Prob (%)	Prob (%)
No DR	1.83	0
MDR increasing	15.02	0
MDR decreasing	2.51	0
NMDR U	0.08	0
NMDR inverted-U	78.79	100
NMDR complex	1.77	0
Total	100	100

Table A.9: From Rubin et al. (2017) (Figure 1b). Endpoint body weight in female exposed perinatally and peripubertally, PND35

	Probabilistic methodology Monte Carlo resampling	Probabilistic methodology Latin-Hypercube resampling
Type of dose-response	Prob (%)	Prob (%)
No DR	2.09	0
MDR increasing	5.16	0
MDR decreasing	6.95	0
NMDR U	0.46	0
NMDR inverted-U	81.43	100
NMDR complex	3.91	0
Total	100	100

Table A.10: From Rubin et al. (2017) (Figure 6b). Endpoint fat mass in female exposed perinatally and peripubertally, PND141

	Probabilistic methodology Monte Carlo resampling	Probabilistic methodology Latin-Hypercube resampling
Type of dose-response	Prob (%)	Prob (%)
No DR	0.3	0
MDR increasing	7.8	0
MDR decreasing	0.7	0
NMDR U	1.1	0
NMDR inverted-U	83.1	100
NMDR complex	7	0
Total	100	100

Table A.11: From Rubin et al. (2017) (Figure 6b). Endpoint fat mass in female exposed perinatally and peripubertally, PND211

	Probabilistic methodology Monte Carlo resampling	Probabilistic methodology Latin-Hypercube resampling
Type of dose-response	Prob (%)	Prob (%)
No DR	2.2	0
MDR increasing	20	0
MDR decreasing	2.6	0

	Probabilistic methodology Monte Carlo resampling	Probabilistic methodology Latin-Hypercube resampling
NMDR U	1.4	0
NMDR inverted-U	70	100
NMDR complex	3.8	0
Total	100	100

Table A.12: From Rubin et al. (2017) (Figure 6b). Endpoint percent fat in female exposed perinatally and peripubertally, PND141

	Probabilistic methodology Monte Carlo resampling	Probabilistic methodology Latin-Hypercube resampling
Type of dose-response	Prob (%)	Prob (%)
No DR	0.15	0
MDR increasing	6.54	0
MDR decreasing	0.49	0
NMDR U	1.68	0
NMDR inverted-U	72.98	100
NMDR complex	18.16	0
Total	100	100

Table A.13: From Rubin et al. (2017) (Figure 6b). Endpoint percent fat in female exposed perinatally and peripubertally, PND211

	Probabilistic methodology Monte Carlo resampling	Probabilistic methodology Latin-Hypercube resampling
Type of dose-response	Prob (%)	Prob (%)
No DR	0.98	0
MDR increasing	19.58	0.1
MDR decreasing	0.87	0
NMDR U	2.9	0
NMDR inverted-U	66.02	99.9
NMDR complex	9.65	0
Total	100	100

Annex A – Assessment of non-monotonicity reported for BPA

In Beausoleil et al. (2016) (EFSA External Report), BPA is reported as the substance under the EFSA remit with the highest number of *in vivo* data sets (35 data sets) for which the authors report a potential NMDR. BPA was also identified in the targeted literature search conducted for this assessment for updating the information. One characteristic of these studies are indications of NMDR present at relatively low-dose BPA exposure, which have been reported for several non-apical endpoints (Lagarde et al., 2015). One limitation of many of these studies is the use of two or three dose groups (in addition to controls), which is not well suited to assess the presence of NMDR with any reasonable certainty.

The presence of NMDRs has also been suggested in several publications based on data from the Consortium Linking Academic and Regulatory Insights on BPA Toxicity (CLARITY-BPA) Program conducted with a wide range of BPA doses⁹ by the US National Toxicology Program in accordance with OECD guidelines. The studies linked to the CLARITY-BPA Program covered previously reported endpoints of potential relevance in the scientific literature, and a large number of effects were measured by different research groups. The participant laboratories received blinded samples, meaning they did not know whether samples had been dosed with BPA or how much, to minimise the potential for bias. Consequently, these studies were considered particularly relevant for addressing NMDR reported for BPA, and were added to those retrieved in the literature search.

This annex covers exclusively the evaluation of the reliability of the NMDR reported by several publication identified by Beausoleil et al. (2016) and the targeted search done for this assessment (including publications from the CLARITY-BPA). One aim of this exercise is to support the EFSA risk assessment on BPA by the CEP Panel.

⁹ Five doses and control.

Table A.1: Studies on BPA with data sets on NMDR included in Beausoleil et al. (2016) (EFSA External Report), and additional studies including those from the BPA-Clarity program assessed for NMDR

Publication, chemical and measured effects	Dose range, No. 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 Beausoleil et al. (2016)							
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 an 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) , <i>in vivo</i> and <i>in vitro</i> effects of BPA, 17β-estradiol (E2) and their mixture on cerebral signally in rats. 1) Extracellular kinase signalling in cerebellar cortex: pERK-IRCellAtP10	Intracerebellar injection of 3 μL per animal of BPA concentrations 10 ⁻¹² –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 <i>in vitro</i> study on primary cerebellar granule cells, range 10 ⁻¹² –10 ⁻⁴ M, N = 5, reported ∩ shape response for induction of ERK phosphorylation.

Publication, chemical and measured effects	Dose range, No. 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
Angle et al. (2013) , effects of in utero BPA exposure in mice 1) Gonadal fat pad weight 2) Renal fat pad weight 3) Serum adiponectin	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 ∩ (CP-2 and CP-3)	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) 3) P _{NMDR} 35 (U)	Some departure from monotonicity seems present but random fluctuation in response also plausible. Beausoleil et al. (2016) also include data set for other endpoints, fulfilling 3 or less checkpoints.
Kendig et al. (2012) , oestrogen-like effects of in utero BPA or 17 α -ethinylestradiol (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 three doses.
Studies not included in Beausoleil et al. (2016) (EFSA External 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)	0.025–50 mg/kg bw day by gavage N = 4	1) Unclear 2) Yes U	1) Intermediate 2) Apical	1) Yes 2) Yes	1) Yes 2) Yes	1) Low for NMDR (MDR for Latin Hypercube) 2) P _{NMDR} 58 (U)	1) Modest effect (less than 20% reduction vs. control) and the probabilistic assessment indicates monotonicity. 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

Publication, chemical and measured effects	Dose range, No. 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
Taylor et al. (2018) , effects of BPA prenatal exposure in mice 1) Gonadal fat pads weight (Figure 1B)	0.005–0.5 mg/kg bw per day by gavage N = 2	1) Yes \cap 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, 1 Sperm DNA methylation (Figure 2)	0.0025–250 mg/kg bw day by gavage N = 6	1) Yes \cap	1) Early effect	1) Yes	1) Unclear as no effects are observed on semen quality in the clarity study (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 year in stop arm for females (Table 4, Figures 1 and 2) 2) Total bile acids at 1 year 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 \cap 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.

Publication, chemical and measured effects	Dose range, No. 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
<p>Uchtmann et al. (2020) (Clarity), effects of BPA early gestation exposure in rats.</p> <p>1) Body weight (Figure 4), 2) Fetal urogenital sinus epithelium thickness (Figure 7)</p>	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.
<p>Li et al. (2019) (literature search), effects of BPA, peri/post pubertal exposure in male mice</p> <p>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)</p>	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 adversity	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.

Publication, chemical and measured effects	Dose range, No. 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
<p>Rubin et al. (2017) (literature search), effects of BPA, perinatal or perinatal and peripubertal exposure in mice</p> <p>1) Body weight in female exposed perinatally and peripubertally (Figure 1)</p> <p>2) Body composition in female exposed perinatally and peripubertally (Figure 6)</p>	0.00025–0.250 mg/kg bw -per day subcutaneous exposure perinatally and by drinking water peripubertally N = 4	1) No 2) No	1) Apical 2) All Intermediate			1) P _{NMDR} 79 for PND 28 and 81 for PND 35 (n). No effects at other PND. 2) P _{NMDR} 83 for fat (g) and 66 for fat (%) (n).	The probabilistic assessment indicates NMDR with ~ 80% probability at PND 28 and 35. However, no sign of NMDR was observed at later ages which somewhat reduces the relevance of this finding in terms of assessing risk. The role of chance for this observation also seems plausible. Same conclusions apply to body composition in females (% fat).
<p>Yang et al. (2016). (literature search), effects of BPA, pubertal exposure in mice</p> <p>1) Body weight (Figure 1)</p> <p>2) Fat mass (Figure 1)</p> <p>3) iWAT and eWAT (Figure 1)</p> <p>4) C/EBP-α (Figure 3)</p> <p>5) SREBP-1c (Figure 3)</p> <p>6) SCD-1 (Figure 3)</p> <p>7) Inflammation (Figure 5)</p> <p>Effects of BPA metabolites on humans</p> <p>1) Plasma Leptin in lean female subjects (Figure 6)</p> <p>2) TNFα levels in lean female subjects (Figure 6)</p>	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).

Publication, chemical and measured effects	Dose range, No. 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
<p>Sharma et al. (2019) (literature search), effects of BPA, exposure in mice</p> <p>1) PPAR (α, β, γ) mRNA</p> <p>2) Protein expression in testes (Figure 2)</p>	4–16 mg/kg per day intraperitoneally N = 3	1) No 2) Yes \cap	1) Early event 2) Early event		Unclear		<p>1) Monotonic decrease in all dose groups; however, controls were lower than the lowest dose group.</p> <p>The apical effect (pattern of histopathological effects) was monotonic.</p>
<p>Zhang et al. (2019) (literature search), human cohort study of pregnant women</p> <p>1) Fasting plasma glucose (Figure 1)</p>	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		<p>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 NMDR for fasting plasma glucose one would expect to see higher risk of GDM at high BPA exposures). As such these findings appear inconsistent.</p>
<p>Zhou et al. (2017) (literature search), effects on BPA, pubertal exposure in male mice (n = 8, 8 week exposure)</p> <p>1) Neuron quantity in the CA3 region of the hippocampus (Figure 4)</p>	0.0005–5 mg/kg bw per day by gavage N = 3	1) No	1) Intermediate				<p>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.</p>

*: CP = checkpoint as defined in Beausoleil et al. (2016) (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 an 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.

‡: The key Monte Carlo resampling results are presented, for additional results see the publications or Appendix A.

In Beausoleil et al. (2016), four studies on BPA were identified and evaluated with respect to the six checkpoints (Tyl et al., 2002; Zsarnovszky et al., 2005; Kendig et al., 2012; 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 four checkpoints or less. Independent review of these studies in Table A.1 is in line with Beausoleil et al. (2016) that the presence of NMDR in these studies 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 Beausoleil et al. (2016) were addressed further by screening for more recent studies that may confirm 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 the opposite direction to the NMDR reported by Kendig et al. (2012). In the more recent CLARITY-BPA study (CLARITY BPA, NTP 2018), no indications of NMDR were observed. Overall findings on NMDR and male fertility appear inconsistent.

Using data from the CLARITY-BPA study, Montevel et al. (2020) reported a non-monotonic dose response between BPA and different morphological features of the mammary gland assessed at postnatal day 21. A summary of these results has been reported in the compilation of academic CLARITY-BPA studies by Heindel et al. (2020). The biological interpretation of some of the outcomes presented in Heindel et al. (2020) (Figures 8 and 9) is not always straightforward (such as the third dimension of a PCA analyses on the 91 structural features). The focus of this discussion is therefore centred on the statistical evaluation of the dose response. The authors assessed non-monotonic dose response (NMDR) using a simple linear step function with the unit function defined as lying between the 25 and 250 $\mu\text{g}/\text{kg}$ bw doses. It is fair to say that use of step functions to assess biological responses is not widespread, but use of such functions is more common in engineering or finance to capture sudden initiated shifts in processes (e.g. sudden changes in voltage or market conditions, respectively). The use of this function in the context of the dose response observed in Figures 8 and 9 appears, based on the authors' description, to be data driven, as the unit function is predefined based on the observed shape of the data. Such an approach is not compatible with formal hypothesis testing. A more conventional approach would be to ask, is there an overall effect? Assuming normal distribution of the outcomes in Figures 8 and 9, a simple F-test shows that only the 'standard deviation of the width in 3D' (Figure 8A) reaches formal significance ($p = 0.03$). The thickness measure (in μm , Figure 8B, $p = 0.06$) was also borderline significant. The effect here appears driven by the 25 $\mu\text{g}/\text{kg}$ bw dose only. The next question would then be to assess whether there is a dose response? Using a more conventional approach, a linear function is non-significant ($p = 0.76$) and the same applies to polynomial of second degree as well ($p = 0.23$). Modelling the data with more biologically based and flexible functions in PROAST (Hill and Exponential) also reaches the same conclusion ($p > 0.05$). The use of PROAST would be considered a more conventional approach for risk assessment and is by no means restrictive as these functions can easily be used to assess non-monotonicity (Badding et al., 2019). The same conclusions were also reached when fitting cubic splines, except when the number of knots reaches the level at which the data are clearly overfitted (see Figure A.1 below).

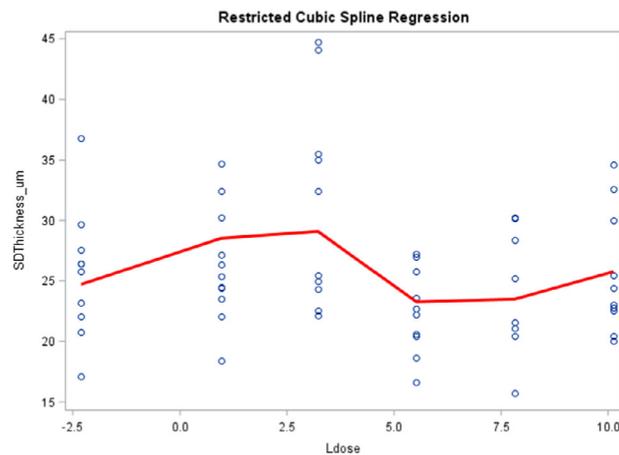


Figure A.1: Restricting cubic spline using four knots as an example of overfitting data to identify NMDR ($p = 0.03$)

Taken together, a reasonable conclusion would be to say that the overall significance identified using ANOVA is driven by the 25 µg/kg bw dose group only (Figure 8A). A formal dose response is, however, not identified by fitting flexible biologically based functions or polynomials that are commonly used to describe biological systems. Without further biological explanations, justifying why the 25 µg/kg group may deviate by other reasons than chance, it is reasonable to conclude that there is no dose response.

If the findings from this paper were replicated in a different setting or if a clear biological explanation explaining the pattern observed in the data were given, then alternative conclusions may very well be reached (based on biology). However, identifying NMDR statistically by fitting flexible functions selected on the basis of how data looks runs the risk of identifying NMDR when random fluctuations and variability in the data are at least equally likely.

Using the six checkpoints, Badding et al. (2019) identified two endpoints (%basophils for females and total bile acids at 1 year in stop arm for males) in the CLARITY-BPA study that fulfilled at least five of the six checkpoints. 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-BPA study where some indications of NMDR at postnatal day 1 have been suggested (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 are a number of *in vivo* studies suggesting NMDR for some early, intermediate and apical effects. The CLARITY-BPA study developed to bridge guideline-compliant research conducted at the FDA with hypothesis-based research investigations conducted by academia on the toxicity of BPA has provided additional evidence.
- Statistical assessments have identified some NMDR data sets extracted from the CLARITY-BPA 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 reporting NMDRs across studies, for several outcomes where NMDR has been suggested, no biological explanation that mechanistically connects the observed 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. *Reproductive Toxicology*, 42, 256–268. <https://doi.org/10.1016/j.reprotox.2013.07.017>
- Badding MA, Barraji L, Williams AL, Scrafford C and Reiss R, 2019. CLARITY-BPA Core Study: analysis for non-monotonic dose-responses and biological relevance. *Food and Chemical Toxicology*, 131, 110554. <https://doi.org/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 2016;13:1027E. <https://doi.org/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. *Toxicology and Applied Pharmacology*, 347, 1–9. <https://doi.org/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. <https://doi.org/10.1111/andr.12176>
- Heindel JJ, Belcher S, Flaws JA, Prins GS, Ho S-M, Mao J, Patisaul HB, Ricke W, Rosenfeld CS, Soto AM, vom Saal FS and Zoeller RT, 2020. Data integration, analysis, and interpretation of eight academic CLARITY-BPA studies. *Reproductive Toxicology*, 98, 29–60. <https://doi.org/10.1016/j.reprotox.2020.05.014>
- Kendig EL, Buesing DR, Christie SM, Cookman CJ, Gear RB, Hugo ER, Kasper SN, Kendzierski JA, Ungi KR, Williams K and Belcher SM, 2012. Estrogen-like disruptive effects of dietary exposure to bisphenol A or 17 α -ethinyl estradiol in CD1 mice. *International Journal of Toxicology*, 31, 537–550. <https://doi.org/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. *Environmental Health Perspectives*, 123, 13. <https://doi.org/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 and Chemical Toxicology*, 133, 110786. <https://doi.org/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. <https://doi.org/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. *Reproductive Toxicology*, 68, 130–144. <https://doi.org/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. <https://doi.org/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. <https://doi.org/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. *Toxicological Science*, 68, 121–146. <https://doi.org/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. <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. <https://doi.org/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. <https://doi.org/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. *Frontiers Endocrinology (Lausanne)*, 10, 262. <https://doi.org/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. <https://doi.org/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. <https://doi.org/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 B – Assessment of non-monotonicity reported for phthalates

Introduction

In Beausoleil et al. (2016) (EFSA External Report), phthalates (DEHP and DBP) are reported within the substances under EFSA remit with the highest number of *in vivo* data sets 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 plasticisers used as FCM under the EFSA domain. Several phthalates are considered as having properties associated with endocrine activity, are classified as toxic for the reproduction (CLP 1B), considered substances of very high concern (SVHC) requiring authorisation 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 per 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, 2019¹⁰).

The reduction of fetal testosterone levels in males is widely recognised as a critical step for the malformation of androgen-dependent reproductive tissues (AOP 288: Collet, 2020) (NAS, 2017¹¹). 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 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 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)¹² *'altered neurodevelopment has been associated with high phthalate exposures in children, as reviewed by*

¹⁰ 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, Wolfie 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>

¹¹ NAS (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. The National Academies Press, Washington, DC. <https://doi.org/10.17226/24758>

¹² 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).

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 effects 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 Beausoleil et al. (2016), complemented with a targeted literature search performed in June 2020 (See Table B.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 selection focused on all experimental evidence on the effects of phthalates on testosterone levels in mammals, including *in vitro* and *ex vivo* studies and/or studies that did not mention NMDR.

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

Table B.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 meh)	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

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 Beausoleil et al. (2016), that was added for completeness. The NMDRs suggested by these publications are summarised in Table B.2.

Table B.2: Studies on phthalates with data sets on NMDR included in Beausoleil et al. (2016) (EFSA External Report)

Publication, chemical and measured effects	Dose range, No. 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 > 100 mg/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) Intermediate	–	–	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 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)

Publication, chemical and measured effects	Dose range, No. 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
<p>Andrade et al. (2006b) (aromatase), effects of DEHP in utero and lactation exposure on aromatase activity at PND 1 & 22 in rats</p> <p>1) Males PND1 2) Females PND1 3) Males PND22 4) Females PND22</p>	0.015–405 mg/kg bw day by gavage N = 10	<p>1) Yes U (All checkpoints met) 2) No (All unmet) 3) No (All unmet) 4) No (Only 3 met)</p>	<p>1) Early event 2) Early event 3) Early event 4) Early event</p>	Yes	No for reproduction (supported by, no effects on apical Repro parameters in the other publications covering this study (Andrade et al. 2006a,c; Grande et al., 2007).	Not analysed	<p>1) Large, overlapping SD and plateau at 4 highest doses.</p> <p>4) Consistent increase except 1 data point</p> <p>Aromatase NMDR could be a possible mechanism for T increases. The role of T increases in possible pathways towards non-reproductive adverse effects is to be investigated.</p>
<p>Grande et al. (2007), effects of DEHP in utero and lactation exposure on reproduction in female rats</p> <p>1) Age vaginal opening 2) Age at 1st oestrus 3) Ano-genital distance PND22 4) Number of nipples at PND13</p>	0.015–405 mg/kg bw day by gavage N = 10	<p>1) No, increase at high doses 2) No, trend for increase at high doses 3) No effect 4) No effect</p>	<p>1) Intermediate 2) Intermediate 3) Apical 4) Intermediate</p>	–	–	Not analysed	<p>Not included in Beausoleil et al. (2016) but added for completeness as reports findings from the same study.</p> <p>Repro parameters not affected (litter size, implantation, birth wt, sex ratio, ano-genital distance at PND22, number of nipples at PND13, ...)</p>

Publication, chemical and measured effects	Dose range, No. 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
<p>Andrade et al. (2006c) (juvenile males), effects of DEHP in utero and lactation exposure on male offspring in rats</p> <p>1) N-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 9) Bw at preputial separation</p>	<p>0.015–405 mg/kg bw day by gavage N = 10</p>	<p>1) No. 2) No. ↑ at 405 only 3) No. ↑ ≥ 5–135, ↓ at 405 4) No effect 5) No effect 6) No. Effects at ≥ 135 mg/kg bw 7) No. No effect 8) No. Trend for delay 9) No</p>	<p>1) Apical 2) Apical 3) Apical 4) Intermediate 5) Intermediate 6) apical effect 7) Apical 8) Apical 9) Apical</p>	–	–	Not analysed	<p>1) Increase of one data point of doubtful biological relevance</p> <p>Not included in Beausoleil et al. (2016) but added for completeness as reports findings from the same study.</p> <p>Authors comment: <i>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 per day</i></p>

Publication, chemical and measured effects	Dose range, No. 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
<p>Christiasen et al. (2010), effects of DEHP in utero and lactation exposure on male reproduction in rats</p> <p>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</p>	<p>3–900 mg/kg bw day by gavage, N = 7</p>	<p>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)</p>	<p>1.–5.; 7.-8.;10. Apical effect 6.;9. early event</p>	<p>–</p>	<p>–</p>	<p>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)</p>	<p>Monotonic effect for ano-genital distance PND1</p>

Publication, chemical and measured effects	Dose range, No. 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
<p>Do et al. (2012), effects of DEHP in utero exposure on male reproduction in mice</p> <p>1) Maternal serum testosterone 2) Fetal male serum testosterone GD18 3) Male offspring testicular testosterone 4) Ano-genital distance PND18 5) Ratio AGD/BW in males 6) Testis weight</p>	0.0005–500 mg/kg bw day, feed once daily (GD9–18), N = 6	<p>1) Yes \cap, trend, (CP-3 and CP-6) 2) No. (Only 3 met) 3) No (only 1 met) 4) No. (Only 2 met) 5) No (Only 1 met) 6) No (All unmet)</p>	<p>1) Intermediate 2) Intermediate 3) Intermediate 4) Apical 5) Intermediate 6) Apical effect</p>	1) ?	1) No for reproductive effects. No effect on litter size	<p>1) P_{NMDR} 47 (\cap) 2) P_{NMDR} 40.24 (\cap) 3) P_{NMDR} 54 (\cap) 4) P_{NMDR} 38.5 (\cap) 5) P_{NMDR} 31.5 (\cap) 6) Low for NMDR (P_{MDR} 43.9)</p>	<p>2) Relative to controls, similar increase except at 500, confirmation by probability assessment 3) No differences vs. control 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 Relevance for phthalate risk assessment is uncertain.</p>
<p>Grande et al. (2006) (juvenile females) effects of DEHP in utero and lactation exposure on female offspring in rats</p> <p>1) Body weight of offspring at PND1 2) Body weight at vaginal opening 3) Kidney weight of dams 4) Body weight at first oestrus</p>	0.015–405 mg/kg bw Gavage, (GD6-PND21), N = 11	<p>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)</p>	<p>1) Apical effect 2) Apical effect 3) Apical effect 4) Apical effect</p>	–	–	Not analysed	<p>No effects observed for E2, progesterone or oestrus cycling. Effects at the highest dose for vaginal and uterine luminal cell height and No. of ovarian atretic tertiary follicles.</p>

Publication, chemical and measured effects	Dose range, No. 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
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	1.5–10,000 mg/kg feed (F3 only up to 7,500 mg/kg feed) (0.1–500 mg/kg bw day) N = 8 (1.5 mg/kg feed in controls)	1 & 2) No. increased incidence at $\geq 7,500$ mg/kg feed 3) No. F3 pregnancy index \downarrow at 7,500 mg/kg feed (359 mg/kg bw) F1 at 10,000 mg/kg feed (543 mg/kg bw) did not produce F2	1-3) Apical effect	–	–	Not analysed	In controls, background exposure was measured. <i>(Remark: included to support the Grande/Andrade studies and to highlight the lack of measurement of background exposure in probably all the other studies)</i>

*: CP = checkpoint as defined in Beausoleil et al. (2016) (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 an 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.

††: The key Monte Carlo resampling results are presented, for additional results, see the publications.

A potential NMDR was observed for maternal serum testosterone in Do et al. (2012), with increases at intermediate doses but less clear NMDR assessment for fetal male serum testosterone at GD18 and male offspring testicular testosterone. Andrade et al. (2006a) observed increased serum testosterone levels in adult male rats exposed during gestation and lactation to 0.045, 0.405 and 405 mg DEHP/kg per 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 B.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 B.3: Selected publications retrieved from the literature search complemented with the references and citations of the retrieved publications

Author	NMDR reported 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 Beausoleil et al. (2016) (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 MCP (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 per 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 modelling 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 N = 4 (0, 250, 500, and 750 mg/kg per day) From PND21 to PND51 by gavage	Few and too high doses
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 modelling 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 behaviour	DEHP exposure on CD-1 mice	Only 3 doses, all below the NOAEL for reproductive effects, the reported NMDR cannot be assessed

Author	NMDR reported by authors	Type of study	Comments on non-monotonicity and analysis
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 Beausoleil et al. (2016) (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 levelling 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 per 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?) <i>'low doses of DEHP (e.g. 10 mg/kg body weight) may stimulate androgen production'</i> In the same study also <i>in vitro</i> 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 suggest NMDR for some phthalate metabolites (e.g. MCP). 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 NMDR but of relevance
James-Todd et al. (2012)	Diabetes	Epidemiological study on phthalates(women)	The suspected NMDR (MnBP and Σ 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)	NMDRs proposed due to non-linear associations (MEHP and cx-MEPP in urine), but data not presented.

Author	NMDR reported by authors	Type of study	Comments on non-monotonicity and analysis
Lee et al. (2004)	Pituitary weight and endocrine alterations	DBP exposure on Sprague Dawley rats N = 5 (0, 20, 200, 2,000 and 10,000 ppm) GD15 to PND21 by diet	N = 5 but MTD
Lind and Lind (2011), 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 per bw) For 13 weeks, by diet	Not enough doses to establish a NMDR but ' <i>Further low-dose investigations are needed to assess non-monotonic dose responses</i> '.
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, 5,000 µg/kg bw per 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 NMDR 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 per day) 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)*	Behavioural effects	DEHP exposure on ICR mice	Inverted-U shape for social play and investigation times in pubertal males, led by a single dose (50 mg/kg per day)

Author	NMDR reported by authors	Type of study	Comments on non-monotonicity and analysis
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 per day) NMDR U shape for oestrogen 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.

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

Evidence on the increase in testosterone observed after phthalates exposure

According to the proposed methodology, for intermediate events such as testosterone levels, the assessment includes different and consecutive steps. First, the available evidence suggesting an NMDR, and the biological plausibility for nonmonotonicity are assessed. Then, if NMDR seems plausible, the biological relevance of the observed effects should be addressed in a subsequent steps. Following this proposal, the evidence on T increases following phthalates exposure and its biological plausibility are described first, presenting all available studies independently of the species, gender or exposure window. Differences in the role of testosterone, among species, sex and developmental/physiological status at the time of exposure or the time of the observations, are considered in the second step, the biological relevance.

A combination of statistical, probabilistic and biological relevance assessments used to analyse the data set from Beausoleil et al. (2016) resulted in the identification of a potentially relevant NMDR observed for serum testosterone by Do et al. (2012), who exposed via micropipette pregnant CD-1 mice with 0, 0.0005, 0.001, 0.005, 0.5, 50 and 500 mg/kg per day DEHP from gestation day (GD) 9 to 18 and examined mothers and male fetuses on GD 18. In particular, an inverted U-shaped dose-response curve was observed for maternal serum testosterone, characterised by a monotonic increase from 0 to 0.005 mg/kg per day and a monotonic decrease from 0.5 to 500 mg/kg per day.

A similar study also included in the data set from Beausoleil et al. (2016) was conducted by Andrade et al. (2006a,b), who exposed by gavage pregnant Wistar rats with DEHP (0, 0.015, 0.045, 0.135, 0.405, 1.215, 5, 15, 45, 135 and 405 mg/kg per day) from GD 6 to post-natal day (PND) 21 and examined offspring rats. Andrade et al. (2006a) observed an increase in serum testosterone levels in adult male offspring (PND 144) exposed at 0.045, 0.405 and 405 mgDEHP/kg per day. In the same study (Andrade et al., 2006b), aromatase activity was determined in hypothalamic/preoptic area brain sections from male and female offspring on PNDs 1 and 22. In males on PND 1, aromatase activity was inhibited at low doses (0.135–0.405 mg/kg per day) and increased at high doses (15, 45 and 405 mg/kg per day). At PND 22, aromatase activity was more affected in females, where an increase in activity was observed at all doses except for 0.045 and 5 mg DEHP/kg per day. Similar observation was reported by Bao et al. (2011), who exposed male Sprague–Dawley rats to DBP (0.1, 1.0, 10, 100, 500 mg/kg bw per day) by gavage during puberty (PND 35–65). Non-statistically significant increase ($+ \approx 30\%$) of serum testosterone level was observed at low dose (1 mg/kg), accompanied by FSH stimulation and non-statistically significant increase of LH and E2.

The literature search performed as described in Table B.1 resulted in several studies reporting NMDR for serum testosterone in different exposure windows and testing conditions. In general, the authors reported unexpected increases in testosterone levels in response to low-dose DEHP (or its main metabolite monoethylhexylphthalate (MEHP)), in contrast to reduced testosterone levels typically associated with exposure to high-dose phthalates. When considering these findings, it should be kept firmly in mind that the timing of DEHP/MEHP exposure, especially whether exposure occurs prenatally

(when negative feedback control on LH secretion is largely inoperative) or postnatally (when negative feedback control on LH secretion is operative) may be critical in determining the biological relevance.

While most available studies have identified either a monotonic decrease of testosterone levels or no clear dose-related effects (depending on the species, tissue, and exposure and observation timelines) following phthalate exposure, there are examples of study authors reporting a potential NMDR for testosterone levels, including experimental studies in rats and mice, as well as epidemiological studies.

- Rat *in vivo* studies:
 - Oudir et al. (2018) observed an increase of serum testosterone level accompanied by Leydig cells hyperplasia after exposure to low doses of DEHP (0.05 mg/kg bw per day). In this study, the authors evaluated effects on reproduction of male Wistar rats after DEHP administration (0.0005, 0.05, 5 mg/kg bw per day) by gavage from PND 21 to 120.
 - Ge et al. (2007) also observed a biphasic effect for serum testosterone level after treatment of male Long-Evans rat pups with DEHP (0, 10, 500, or 750 mg/kg bw) from PND 21 to 49. In particular, the authors observed an increase at low-dose DEHP (10 mg/kg) and a decrease at high-dose DEHP (750 mg/kg), with concomitant advancement in puberty onset at the low dose. In the same study the *in vitro* investigation of the direct action of MEHP on Leydig cell steroidogenic capacity resulted in an increase of LH-stimulated testosterone production at low MEHP concentration (100 µmol/L) and inhibition at 10 mM.
- Mice *in vivo* studies
 - An increase in circulating testosterone was also observed at 2.6 mg/kg per day by Repouskou et al. (2019), who reported a concomitant increase in estradiol and LH levels in adult male offspring (PND 90) after exposure of pregnant C57/BL6 mice by diet from GD 0.5 to delivery to a mixture of phthalates identified as relevant in a human study (MBP, MBZP, MEHP and MINP) (0.26, 2.6, 13 mg/kg per day).
- Human epidemiological studies
 - Epidemiological studies based on urinary phthalate metabolites also observed NMDR for testosterone level and genes involved in the steroidogenesis pathway (Adibi et al., 2010), and suggests potential relationships with aromatase activity (Meeker et al., 2009) and possible alteration in HPG feedback (Pan et al., 2011).

In addition, other experimental studies evaluating the effects of phthalates observed a stimulatory effect of phthalate exposure on testosterone concentration. These studies were not retrieved in the search described in Table B.1 either because did not describe the observed effect as 'non-monotonic dose-response', and/or because were performed *in vitro* or *ex vivo*. Table B.2 summarises the experimental studies providing evidence for an increase in testosterone levels in response to phthalate 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 in postnatal animals, which are opposite to the generally observed reduction in testosterone levels at high doses in postnatal animals. In contrast, when there is fetal exposure of rats to phthalates, the effects on testosterone are generally monotonic with higher doses suppressing testosterone, which is linked to adverse reproductive effects by the CEP Panel and others (e.g. NAS 2017 meta-analysis). Evidence that fetal exposure to 'low dose' phthalates can increase testosterone levels has been reported in the Lin et al., 2008 study, but not confirmed by other studies at similar doses (Mahood et al., 2007; Howdeshell et al., 2008; Gray et al., 2009).

Table B.4: Summary of experimental studies proving evidence for increases in testosterone levels in response to phthalates' exposure

Publication ¹ , chemical, and measured effects	Evidence for testosterone (T) increase	Additional effects observed at the same dose
Do et al. (2012), gestational exposure (GD 9–18) via micropipette of pregnant CD-1 mice with DEHP (0, 0.0005, 0.001, 0.005, 0.5, 50 and 500 mg/kg per day). Effects on reproduction on mothers and male foetus were analysed at GD18.	NMDR observed for maternal serum testosterone at GD18, with ↑ from 0 to 0.005 mg/kg per day and ↓ from 0.5 to 500 mg/kg per day	Fetal male serum testosterone ↑ at G18 from 0.0005 to 50 mg/kg/day and slight ↓ at 500 mg/kg/day Slight ↑ of AGD at G18 from 0.0005 to 0.5 mg/kg/day with clear ↑ at 0.005 mg/kg/day
Repouskou et al. (2019), gestational exposure (GD 0.5 to delivery) by diet of pregnant C57/BL6 mice with a mixture of phthalates (MBP, MBzP, MEHP and MINP) (0, 0.26, 2.6, 13 mg/kg per day). Effects on reproduction on offspring analysed at PND 1, 21, 90.	↑ circulating testosterone in adult male offspring (PND90) at 2.6 mg/kg per day	Estradiol and LH levels ↑ in adult male offspring (PND90) at 2.6 mg/kg per day Cyp19a1 gene expression ↓ in developing female ovaries of the 2.6 mg/kg per day at PND21, and non-statistically significant ↓ at PND90. Cyp17a1 gene expression ↓ at the same dose at PND90. AGD ↓ at PND21 in male and female offspring at 2.6 mg/kg per day Gonad weight/BW ↓ in male and female offspring of the 2.6 mg/kg per day at PND21 and in adult females at PND90. At the same dose, BW ↓ was also observed in female offspring on PND21 Seminiferous tubules aberrations ↑ in male gonad of the 2.6 mg/kg per day at PND21 Reduced number of secondary follicles and increased atresia in female ovaries of the 2.6 mg/kg per day at PND21. Similar effect observed at PND90, where also primary follicles were affected.
Lin et al. (2008), gestational exposure (GD2 to GD20) by gavage of pregnant Long-Evans rats with DEHP (0, 10, 100, 750 mg/kg per day). Effects on reproduction on male offspring analysed at GD21.	Biphasic effect observed on fetal Leydig cells (FLC) testicular testosterone at GD21 in male offspring (+50% at 10; –66% at 750).	Insulin-like growth factor I (Igf1) and Kit ligand (Kitl) ↑ at 10 mg/kg Increase in number and frequency of FLCs clusters at 10 mg/kg per day Note: no confirmed by other studies (see references in the text)
Andrade et al. (2006 a,b), gestational and perinatal exposure (GD 6 to PND 21) by gavage of pregnant Wistar rats and offspring with DEHP (0, 0.015, 0.045, 0.135, 0.405, 1.215, 5, 15, 45, 135 and 405 mg/kg per day). Effects on reproduction in adult male offspring analysed at PND144 (Andrade et al., 2006a), and effects on brain aromatase (hypothalamic/preoptic area) in male and female pups analysed at PND 1 and 22 (Andrade et al., 2006b).	Serum testosterone levels ↑ in adult male rats (PND144) exposed at 0.045, 0.405 and 405 mg/kg per day	Reduction in daily sperm production and weight of seminal vesicle with coagulating glands at 405 mg/kg per day in treated adult male rats (PND144). At the same dose level, histopathological changes in testis were also observed. Moreover, in one animal with very high serum testosterone concentration (21.8 ng/mL vs control mean = 3.5 ng/mL) light focal Leydig cell hyperplasia was noticed. Abnormal sperm morphology was reported at 0.045 mg/kg/day (PND144). Brain aromatase activity on PND 1 in males ↓ at 0.405 mg/kg per day and ↑ at 405 mg/kg/day. At PND 22, brain aromatase activity ↑ in males only at 0.405 mg/kg per day, and in females at 0.405 and 405 mg/kg per day.

Publication ¹ , chemical, and measured effects	Evidence for testosterone (T) increase	Additional effects observed at the same dose
Jones et al. (2015), <i>ex vivo</i> effects on steroidogenesis in male Sprague Dawley rat testis at PND3 treated with MEHP (10 µmol/L for 3 days).	Basal testosterone production ↑ in rats testes at PND3 after <i>ex vivo</i> exposure to 10 µmol/L MEHP	Testosterone stimulation observed at 10 µmol/L MEHP was normalized by Genistein cotreatment.
Ge et al. (2007), prepubertal exposure (PND 21 to 49) by gavage of male Long-Evans rats with DEHP (10, 500, or 750 mg/kg bw per day). Effects on reproductive endpoints and puberty onset analysed during and immediately after treatment.	Biphasic effect observed for serum testosterone levels with ↑ at 10 mg/kg and ↓ at 750 mg/kg in male rats at PND 49.	Advanced puberty onset (39.7 days ± 0.1 days vs C = 41.5 days ± 0.1 days), at 10 mg/kg. In contrast, delayed puberty (46.3 days ± 0.6 days) was observed at 750 mg/kg. Seminal vesicle weight and body weight ↑ at 10 mg/kg on PND 49, but not on the day of preputial separation (index of pubertal onset). Non-statistically significant ↓ on pituitary LH b subunit (Lhb) gene expression at 10 mg/kg. Further investigation of MEHP <i>in vitro</i> effects on Leydig cell steroidogenic capacity (obtained from 35-day-old rats) resulted in a biphasic effect with ↑ LH-stimulated testosterone production at 100 µM and ↓ at 10 mM. No effects were observed in cell viability at any dose. Increase in MEHP-induced T production was also observed in absence of LH at 1 mM.
Akingbemi et al. (2004), pre-pubertal, pubertal and post-pubertal exposure (PND 21 to 48, 90, or 120) by gavage of male Long-Evans rats with DEHP (0, 10, or 100 mg/kg/day). Effects on male reproduction analysed at PND 48, 90 or 190.	Serum testosterone ↑ in male rats at 10 and 100 mg/kg/day DEHP treatment from PND 21 to 90, and in the 100 mg/kg/day group exposed from PND 21 to 120.	Serum luteinizing hormone (LH) ↑ at 10 and 100 mg/kg/day DEHP treatment from PND 21 to 90 and 120. Basal and LH-stimulated T production per Leydig cell , measured <i>ex vivo</i> , ↓ at 10 and 100 mg/kg/day DEHP treatment from PND 21 to 90, and in the 100 mg/kg/day group exposed from PND 21 to 120. Cell cycle proteins (PCNA, P53, cyclin D3, cyclin G1) gene expression in Leydig cells (LC) and LC number ↑ at 10 and 100 mg/kg/day DEHP treatment from days 21 to 90. At the same doses, LC number and thymidine incorporation by LC ↑ after treatment from PND 21 to 120. Serum E2 levels, LH-stimulated E2 production by LC, and aromatase gene expression in LC ↑ at 10 and 100 mg/kg/day DEHP treatment from PND 21 to 48. Basal LC E2 production ↑ only at 100 mg/kg/day.
Oudir et al. (2018), pre-pubertal, pubertal and post-pubertal exposure (PND 21 to 120) by gavage of male Wistar rats with DEHP (0.0005, 0.05, 5 mg/kg bw per day). Effects on male reproduction analysed at PND 120.	Serum testosterone level ↑ at 0.05 mg/kg bw per day in male rats.	Absolute epididymis and testis weight , and Sertoli cell number ↓ at 0.05 mg/kg bw per day. At the same dose, moderate oligospermia and LC hyperplasia was also observed.
Kurahashi et al. (2015), pre-pubertal exposure (PND28 to 56 or 84) by inhalation of male Wistar Rats with DEHP (5 or 25 mg/m ³ , 6 h/day). Effects on reproduction were analysed at PND 56 and 84.	Plasma testosterone concentration ↑ at both doses in male rats on PND 56 and 84.	Seminal vesicles weight ↑ at both doses on PND 56 and 84.

Publication ¹ , chemical, and measured effects	Evidence for testosterone (T) increase	Additional effects observed at the same dose
Bao et al. (2011), pubertal exposure (PND35 to 65) by gavage of male Sprague–Dawley rats with DBP (0.1, 1.0, 10, 100 and 500 mg/kg bw per day). Effects on reproduction were analysed after treatment.	Non-statistically significant increase (+≈30%) of serum testosterone level at 1 mg/kg in male rats.	Serum FSH ↑ at 1 mg/kg, and a trend of non-statistically significant increase was also observed for serum E2 and LH . SOD gene expression ↑ in the same dose group.
Ljungvall et al. (2005), prepubertal exposure by intramuscular injection of male boars (6 weeks) to DEHP (50 mg/kg, twice week) for 5 weeks. Effects on reproduction were analysed during, after the treatment and at 7.5 months	Testosterone concentration ↑ in male boars at 7.5 months of age.	Increased area of the Leydig cells in the testicles at 7.5 months of age.
Zhao et al. (2012), <i>ex vivo</i> effects of MEHP (0, 2, 20, 200, 2000 μmol/L for 24 h) on steroidogenesis of different stages of Long-Evans male rat Leydig cells, progenitor (PLCs, PND21), immature (ILCs, PND35) and adult (ALCs, PND49).	Testosterone concentration ↑ in ALC at 20 and 200 μmol/L and ↓ at 2000 μmol/L.	Reactive oxygen species (ROS) ↑ at 20 and 200 μmol/L in ALCs, together with stimulation of steroidogenic acute regulatory (StAR), cytochrome P450 side-chain cleavage (P450_{scc}), 3β-hydroxysteroid dehydrogenases (3β-HSD) and 17β-hydroxysteroid dehydrogenases (17β-HSD). Opposite effect on steroidogenesis observed at 2000 μmol/L.
Savchuk et al. (2015), <i>ex vivo</i> effects of MEHP, MBP, MBeP (1, 3, 10, 30, 90 μmol/L for 48 h) on steroidogenesis of 1-month-old male CBA/Lac, C57BL/6j mouse Leydig cells.	Basal testosterone concentration ↑ on Leydig cells of both mice strains at 90 μmol/L MEHP.	Perturbation in the redox state of Leydig cells of both mice strains treated with 90 μmol/L MEHP. ATP levels ↓ in both mice strains Leydig cells treated with 90 μmol/L MEHP. StAR protein expression ↑ in both mice strains Leydig cells treated with 90 μmol/L MEHP.
Forgacs et al. (2012), <i>in vitro</i> effects of MEHP (100 μmol/L) on steroidogenesis of male mouse BLTK1 Leydig cells.	Basal testosterone concentration ↑ at 100 μmol/L MEHP.	rhCG-induced T levels concentration ↓ at 100 μmol/L MEHP.
Gunnarsson et al. (2008), <i>in vitro</i> effects of MEHP (0, 10, 25, 50, 75, 100 μmol/L for 24 h) on steroidogenesis of male mouse Leydig tumour cell line (MLTC-1) and female mouse granulosa tumour cell line (KK-1).	Monotonic ↑ of basal testosterone concentration at 25–100 μmol/L in MLTC-1.	Monotonic ↑ of basal progesterone concentration at 25–100 μmol/L in MLTC-1, but ↓ hCG-stimulated progesterone synthesis at 100 μmol/L in MLTC-1 cells. Similar effect on progesterone stimulation observed in female KK-1 at 100 μmol/L. Hormone-sensitive lipase (HSL) and 3-hydroxy-3-methylglutaryl coenzyme (HMG-CoA) ↑ at 10–100 μmol/L in MLTC-1.

1: Publications are sorted by exposure window (i.e. gestational perinatal pubertal exposure).

Biological plausibility assessment for testosterone levels NMDR following phthalate exposure

To assess the biological plausibility of possible NMDRs for phthalate effects on testosterone production, it is essential to consider the timing (animal age) of phthalate exposure, the species used and whether the study is in males or females, as the effects in these different situations occur in very different hormone regulatory contexts and with potentially different consequences. For example, high dose phthalate exposure profoundly suppresses steroidogenesis by the fetal rat testis whereas there are no comparable effects in the mouse. Similarly, in the female fetus, ovarian steroidogenesis is quiescent and is therefore not susceptible to phthalate inhibition at this stage. Also important is that in fetal male rodents, the hypothalamo-pituitary-testis regulatory axis is not active until around birth so feedback regulation of testicular steroidogenesis via LH modulation is only possible after birth through to adulthood. Finally, with regard to phthalate-induced inhibition of fetal testis steroidogenesis, it is

important to keep in mind that in rodents, fetal testis steroidogenesis during the critical masculinisation programming window is locally regulated and does not require LH drive (contrary to the situation after birth) whereas in humans and non-human primates, steroidogenesis during the comparable period is completely hCG-dependent, hCG being an LH-like molecule produced by the placenta.

The steroidogenic pathway involves the sequential action of five different enzymes to result in the end product, testosterone (Scott et al., 2009). This androgen can be further metabolised in target tissues to the more potent androgen 5-alpha-dihydrotestosterone (via 5-alpha-reductase enzymes) or to oestradiol (via the aromatase enzyme). In postpubertal females, testosterone or androstenedione produced via one cell type (theca) in the ovary is predominantly converted to estradiol by another cell type (granulosa), such that estradiol rather than testosterone is the normal major product of the adult ovary. In contrast, in the adult testis of most species, including rodents and man, only small amounts of testosterone produced by the Leydig cells is converted locally to oestradiol.

Several reviews confirm the capacity of phthalates to disrupt different steps linked to testosterone synthesis, both in fetal males (e.g. Scott et al., 2009) and in adult female mice (e.g. Hannon and Flaws, 2015). In addition to effects on testosterone, biphasic responses (increase and decrease following exposure to DEHP or the metabolite MEHP) have been reported for progesterone and for the steroidogenic acute regulatory protein "StAR" (Hannon and Flaws, 2015) while monophasic responses (evidence suggest consistently either increase or decrease) have been observed for other steps (e.g. Scott et al., 2009; Hannon and Flaws, 2015). Biphasic responses have been also observed in the metabolic testosterone pathway, including for the successor estradiol and for the main enzyme Cytochrome-P450 aromatase "CYP19A1" (Hannon and Flaws, 2015), the second with confirmed NMDR in males at birth in the Andrade et al. (2006b) study and supported by *in vitro* and *in silico* molecular docking studies confirming the elevated binding affinity of phthalates to CYP19A1 (Gupta et al., 2010, Ahmad et al., 2017). This evidence confirms that phthalates, and in particular DEHP and its metabolite MEHP, may disrupt steroidogenesis at different levels resulting in complex responses, suggesting several possible biologically plausible options for explaining a NMDR in testosterone levels following phthalates exposure, that include biphasic responses during the synthesis or metabolism, as well the combined results of monotonic mechanisms some triggering T increase and others triggering T reductions. It should be noted that due to the variability in the preferred steroidogenesis pathway between species, tissues, cell types or physiological/development status, the observations reported in Table B.2 could be explained though different single or combined mechanisms and the described effects may be context-specific and not generally applicable.

In addition to direct disruption of steroidogenesis pathways, there are a number of feedback mechanisms that could explain a NMDR for testosterone levels following phthalates exposure after birth. In males, the LH mediated HPG feedback mechanism regulates testosterone productions by the Leydig cells during postnatal life. Thus any exposure that reduces testosterone production by the Leydig cells will lead to a compensatory increase in LH secretion from the pituitary to restore normal testosterone levels; this compensatory mechanism is not operative in fetal life during the masculinisation programming window. Studies such as Akingbemi et al. (2004), Lin et al. (2008) or Kurahashi et al. (2015) are compatible with an hypothesised mechanism of action with chronic DEHP exposure triggering a reduction in basal and LH-induced testosterone production, LH increase, and Leydig cell hyperplasia that under certain conditions produces an increase in testosterone levels above normal control levels. The elements linked to the NMDR assessment are described in Figure 1B, the chronic phthalate-induced reduction in testosterone (T) synthesis leads to an initial reduction in T levels that triggers the compensatory feedback mechanism which can include Leydig cell hyperplasia if there is long-term disruption. 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 mechanism results in Leydig cell hyperplasia but with reduced T production capacity per Leydig cell. 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.

LH is not involved in the control of testosterone production in fetal Leydig cells during the masculinisation window in rodents (Scott et al., 2009), and the available evidence suggests that phthalates exert their negative effect on steroidogenic enzymes and on testosterone production in rats by preventing the normal age-dependent removal of a negative regulatory factor (COUP-TFII) that suppresses steroidogenic enzyme expression (van den Driesche et al., 2012). In contrast, in humans

hCG action via the LH receptor is the main driver of steroidogenesis by fetal Leydig cells, which may explain why the available evidence indicates that phthalates that suppress testosterone production by fetal rat Leydig cells have no inhibitory effect on hCG-stimulated testosterone production by fetal human Leydig cells (Kilcoyne & Mitchell 2019). In this regard, one epidemiological study has shown an association between phthalate exposure during late gestation and the expression of the hCG gene in placental trophoblast obtained at birth (Adibi et al., 2015).

In the ovaries, testosterone or androstenedione is synthesised by the theca cells and then diffuses to the granulosa cells to be converted into oestradiol by aromatase (Somboonporn and Davis, 2004). The process is regulated by the HPG axis with feedback mechanisms involving LH and FSH according to the stage of the ovarian cycle (Hannon and Flaws, 2015); consequently increases in T levels in females through an overcompensation of the feedback mechanisms is also a plausible hypothesis for the increases observed in females.

To add complexity, it should be noted that phthalates also alter other hormones and proteins within the steroidogenesis pathways (Scott et al., 2009; Hannon and Flaws, 2015; Hlisnikova et al., 2020); therefore, T levels could be also indirectly affected through alterations of levels of precursors in the delta-5 or delta-4 steroidogenesis pathways.

In addition, and specifically for humans, circulating testosterone is mostly bound to albumin and Sex Hormone-Binding Protein (SHBG), and several phthalates including DEHP have high binding affinity for SHBG (Sheikh et al., 2016); therefore, another possible mechanism is associated to alterations in binding capacity and free-bound testosterone levels.

In conclusion, a number of complementary mechanisms could explain an NMDR for testosterone associated to postnatal effects of phthalates exposure; not all mechanisms are relevant for all species/sex, cell types, tissues and developmental/physiological situations; but in conjunction, provide a reasonable plausibility covering males and females in rodents and humans but this would only apply where there is postnatal exposure.

None of the compensatory mechanisms outlined are functional in fetal life during the masculinisation programming window. This is in line with the assessment of the evidence summarised in Table B.4, showing indications of increase in testosterone levels associated with 'low dose' phthalates exposure for postnatal animals, while very limited and inconsistent for the fetal situation.

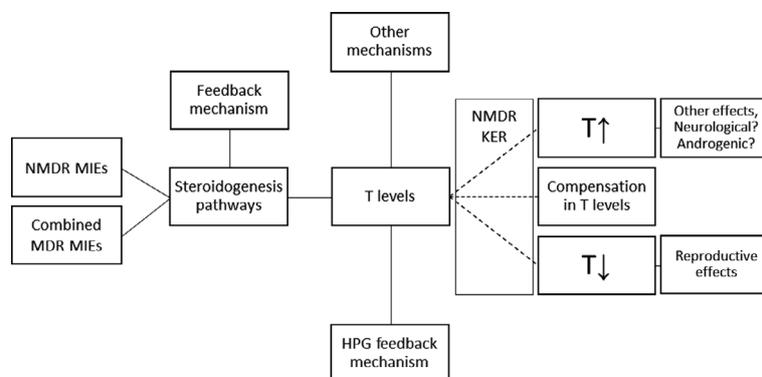


Figure B.1: Hypothesis of an AOP-based mechanistic understanding of the inverted U-shaped curve for testosterone level as intermediate event. T (testosterone), MDR (monotonic dose-response), MIE (molecular initiating event), KER (key event relationship)

Possible role of T increase in the pathways towards adversity

The most recent EFSA risk assessment on phthalates (EFSA CEP Panel, 2019) confirmed the selection of reproductive effects (with the testis as target organ) for setting the Reference Point for establishing the HBGV for DEHP, DBP and BBP. The opinion also confirmed the role of the reduction of the fetal testosterone production in rats as a critical step in the reproductive toxicity of these phthalates. In this review, no consistent evidence of NMDR has been retrieved for this exposure period; and based of the available information is concluded that these effects are associated with a dose-related monotonic reduction in fetal testosterone levels in male fetuses during the masculinisation window.

In addition to the reproductive effects in males, the CEP Panel opinion and other reviews highlighted uncertainties related to the lack of sufficient investigation of other non-reproductive effects,

in particular neurological effects and effects on the immune and metabolic systems. For some of these effects, NMDR and links with steroidogenesis and hormonal control have been proposed by some authors.

This report has identified some indications of NMDR for DEHP postnatal effects, which are biologically plausible. According to the proposed approach, as testosterone levels are an intermediate event, the next steps should be to assess the possible biological relevance of these effects, in particular if a (quantitative) relation between these effects and an adverse outcome (i.e. apical effect) can be established, ideally through a mechanistic sequence (AOP).

There is information indicating that postnatal increases in testosterone levels under certain conditions may trigger pathways resulting in adverse outcomes. Examples cover 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 with overexpression of androgens (Hotchkiss et al., 2007; Martin et al., 1998). There are also, some epidemiological studies linking 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). As mentioned, a comprehensive in depth assessment of these effects has not been performed yet, and is outside the scope of this Opinion.

In addition to the postnatal testosterone levels, a related and plausible NMDR has been identified for aromatase. In humans, brain aromatase has been associated with personality traits and neurobehavioral disorders (Takahashi et al., 2018; Sarachana et al., 2011), providing an additional mechanistic link.

In summary, there are some indications associating increased testosterone levels with non-reproductive apical effects similar or associated with those described as not fully confirmed but of possible concern in the CEP Panel opinion. The establishment of a quantitative association with the related adverse apical outcomes would require a full assessment and is outside this mandate. Similarly to the association of testosterone reduction with reproductive effects, the possible adverse consequences of postnatal testosterone increases would depend on the specific exposure window, sex, species and other factors. In this regard, it should be noted that rodent studies that have involved experimental increase in exposure to androgens in fetal life have shown there is no consequence for exposed males whereas this exposure induces abnormal masculinisation of females (Dean et al., 2012). The aetiology of the hypothesised effects is mostly multifactorial and still poorly understood. The final effect may also depend on the effective internal concentration of MEHP, which seems to be partly responsible for the effect of DEHP.

Answer to the questions (proposed approach)

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

- There is some evidence supporting NMDR for postnatal testosterone levels, covering several species and developmental/physiological status (see Table B.4 for details). In contrast, in utero exposure is linked to a monotonic decrease of fetal testosterone levels in males.

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 the plausibility of the observed NMDR for postnatal testosterone as intermediate effect. Possible mechanisms include NMDR in the T synthesis or metabolisms (as NMDR has been reported for aromatase); a combination of monotonic responses on different steps of the steroidogenesis pathway; disturbance of the HPG feedback mechanism; or other indirect mechanisms.
- Some non-reproductive effects have been reported for DEHP and other phthalates. A comprehensive assessment of these effects is pending, but some indications of NMDR and links with steroidogenesis and hormonal control have been suggested.
- There are experimental and epidemiological studies that link the increase of testosterone with effects similar or linked to those reported for phthalates and overexpression of androgens.
- The establishment of a quantitative relationship between the increase in T levels and the observed effects should consider the multifactorial aetiology 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 postnatal testosterone levels, that could be linked to different routes to adverse effects.
- The effects observed deserve further consideration; however, 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. <https://doi.org/10.1289/ehp.0900788>
- Adibi JJ, Lee MK, Naimi AI, Barrett E, Nguyen RH, Sathyanarayana S, Zhao Y, Thiet M-P, Redmon JB and Swan SH, 2015. Human chorionic gonadotropin partially mediates phthalate association with male and female anogenital distance. *The Journal of Clinical Endocrinology and Metabolism*, 100, E1216–E1224. <https://doi.org/10.1210/jc.2015-2370>
- 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. <https://doi.org/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. <https://doi.org/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. <https://doi.org/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. <https://doi.org/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. <https://doi.org/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 sharp-tooth catfish (*Clarias gariepinus*) exposed to waterborne di-(2-ethylhexyl) phthalate. *General and Comparative Endocrinology*, 254, 22–37. <https://doi.org/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. <https://doi.org/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. <https://doi.org/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. *Environmental Research*, 172, 194–201. <https://doi.org/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. <https://doi.org/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 japonica*) testis following pre-pubertal exposure to di(n-butyl) phthalate (DBP). *Comparative Biochemistry and Physiology C-Toxicology & Pharmacology*, 166, 24–33. <https://doi.org/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. <https://doi.org/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. <https://doi.org/10.1093/toxsci/kfq147>

- Botelho GKG, 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. <https://doi.org/10.1007/s00244-009-9317-8>
- Braun JM, 2017. Early-life exposure to EDCs: role in childhood obesity and neurodevelopment. *Nature Review Endocrinology*, 13, 161–173. <https://doi.org/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. <https://doi.org/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. <https://doi.org/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. <https://doi.org/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. *Reproductive Toxicology*, 30, 313–321. <https://doi.org/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. <https://doi.org/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 Behaviour*, 71, 41–48. <https://doi.org/10.1016/j.yhbeh.2015.03.008>
- Dean A, Smith LB, Macpherson S and Sharpe RM, 2012. The effect of dihydrotestosterone exposure during or prior to the masculinization programming window on reproductive development in male and female rats. *International Journal of Andrology*, 35, 330–339. <https://doi.org/10.1111/j.1365-2605.2011.01236.x>
- 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. <https://doi.org/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. <https://doi.org/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. *Reproductive Toxicology*, 34, 614–621. <https://doi.org/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. <https://doi.org/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, Coconcelli 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. <https://doi.org/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. <https://doi.org/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. *Archives of Toxicology*, 92, 501–512. <https://doi.org/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. <https://doi.org/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. <https://doi.org/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. <https://doi.org/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. <https://doi.org/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. <https://doi.org/10.1016/j.tox.2006.10.005>
- Gray LE, Jr, Barlow NJ, Howdeshell KL, Ostby JS, Furr JR and Gray CL, 2009. Transgenerational effects of di (2-ethylhexyl) phthalate in the male CRL:CD(SD) rat: added value of assessing multiple offspring per litter. *Toxicological Sciences*, 110, 411–425. <https://doi.org/10.1093/toxsci/kfp109>
- 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. <https://doi.org/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. <https://doi.org/10.1095/biolreprod.115.129148>
- Hannon PR and Flaws JA, 2015. The effects of phthalates on the ovary. *Frontiers in Endocrinology*, 6. <https://doi.org/10.3389/fendo.2015.00008>
- 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. <https://doi.org/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. *Physiological Behavior*, 207, 7–14. <https://doi.org/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.
- Hliseníková H, Petrovičová I, Kolena B, Šidlovská M, Sirotkin, A. Effects and mechanisms of phthalates' action on reproductive processes and reproductive health: a literature review. *International Journal of Environmental Research and Public Health*, 2020, 17, 6811. <https://doi.org/10.3390/ijerph17186811>
- Hotchkiss AK, Lambright CS, Ostby JS, Parks-Saldutti L, Vandenberg 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. <https://doi.org/10.1093/toxsci/kfm002>
- Howdeshell KL, Wilson VS, Furr J, Lambright CR, Rider CV, Blystone CR, Hotchkiss AK and Gray Jr LE, 2008. A mixture of five phthalate esters inhibits fetal testicular testosterone production in the Sprague-Dawley rat in a cumulative, dose-additive manner. *Toxicological Sciences*, 105, 153–165. <https://doi.org/10.1093/toxsci/kfn077>
- Hu JMY, Arbuckle TE, Janssen P, Lanphear BP, Braun JM, Platt RW, Chen A, Fraser WD and McCandless LC, 2020. Associations of prenatal urinary phthalate exposure with preterm birth: the Maternal-Infant Research on Environmental Chemicals (MIREC) Study. *Canadian Journal of Public Health*, 111, 333–341. <https://doi.org/10.17269/s41997-020-00322-5>
- Huang YC, Sun FJ, Tan HL, Deng YF, Sun ZQ, Chen HX, Li J and Chen D, 2019. DEHP and DINP induce tissue- and gender-specific disturbances in fatty acid and lipidomic profiles in neonatal mice: a comparative study. *Environmental Science & Technology*, 53, 12812–12822. <https://doi.org/10.1021/acs.est.9b04369>
- James-Todd T, Stahlhut R, Meeker JD, Powell SG, Hauser R, Huang TY and Rich-Edwards J, 2012. urinary phthalate metabolite concentrations and diabetes among women in the National Health and Nutrition Examination Survey (NHANES) 2001–2008. *Environmental Health Perspectives*, 120, 1307–1313. <https://doi.org/10.1289/ehp.1104717>

- Jones S, Boisvert A, Francois S, Zhang L and Culty M, 2015. In utero exposure to di-(2-ethylhexyl) phthalate induces testicular effects in neonatal rats that are antagonized by genistein cotreatment. *Biology of Reproduction*, 93, 92. <https://doi.org/10.1095/biolreprod.115.129098>
- Kasper-Sonnenberg M, Wittsiepe J, Wald K, Koch HM and Wilhelm M, 2017. Pre-pubertal exposure with phthalates and bisphenol A and pubertal development. *PLOS One*, 12. <https://doi.org/10.1371/journal.pone.0187922>
- Kilcoyne KR and Mitchell RT, 2019. Effect of environmental and pharmaceutical exposures on fetal testis development and function: a systematic review of human experimental data. *Human Reproduction Update*, 25, 397–421. <https://doi.org/10.1093/humupd/dmz004>
- Kurahashi N, Kondo T, Omura M, Umemura T, Ma MY and Kishi R, 2005. The effects of subacute inhalation of Di (2-ethylhexyl) phthalate (DEHP) on the testes of prepubertal Wistar rats. *Journal of Occupational Health*, 47, 437–444. <https://doi.org/10.1539/joh.47.437>
- Lee KY, Shibutani M, Takagi H, Kato N, Takigami S, Uneyama C and Hirose M, 2004. Diverse developmental toxicity of di-n-butyl phthalate in both sexes of rat offspring after maternal exposure during the period from late gestation through lactation. *Toxicology*, 203, 221–238. <https://doi.org/10.1016/j.tox.2004.06.013>
- Lehmann KP, Phillips S, Sar M, Foster PM and Gaido KW, 2004. Dose-dependent alterations in gene expression and testosterone synthesis in the fetal testes of male rats exposed to di (n-butyl) phthalate. *Toxicological Sciences*, 81, 60–68. <https://doi.org/10.1093/toxsci/kfh169>
- Lin H, Ge RS, Chen GR, Hu GX, Dong L, Lian QQ, Hardy DO, Sottas CM, Li XK and Hardy MP, 2008. Involvement of testicular growth factors in fetal Leydig cell aggregation after exposure to phthalate in utero. *Proceedings of the National Academy of Sciences of the United States of America*, 105, 7218–7222. <https://doi.org/10.1073/pnas.0709260105>
- Lind PM and Lind L, 2011. Circulating levels of bisphenol A and phthalates are related to carotid atherosclerosis in the elderly. *Atherosclerosis*, 218, 207–213. <https://doi.org/10.1016/j.atherosclerosis.2011.05.001>
- Ljungvall K, Karlsson P, Hulten F, Madej A, Norrgren L, Einarsson S, Rodriguez-Martinez H and Magnusson U, 2005. Delayed effects on plasma concentration of testosterone and testicular morphology by intramuscular low-dose di(2-ethylhexyl)phthalate or oestradiol benzoate in the prepubertal boar. *Theriogenology*, 64, 1170–1184. <https://doi.org/10.1016/j.theriogenology.2005.02.003>
- Mahood IK, Scott HM, Brown R, Hallmark N, Walker M and Sharpe RM, 2007. In Utero Exposure to Di (n-butyl) phthalate and testicular dysgenesis: comparison of fetal and adult end points and their dose sensitivity. *Environmental Health Perspectives*, 115, 55–61. <https://doi.org/10.1289/ehp.9366>
- 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. <https://doi.org/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. <https://doi.org/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. <https://doi.org/10.2164/jandrol.108.06403>
- 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, 396–402. <https://doi.org/10.1289/ehp.1103582>
- Miodovnik A, Edwards A, Bellinger DC and Hauser R, 2014. Developmental neurotoxicity of ortho-phthalate diesters: review of human and experimental evidence. *Neurotoxicology*, 41, 112–122. <https://doi.org/10.1016/j.neuro.2014.01.007>
- 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. <https://doi.org/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. *Reproductive Toxicology*, 75, 33–39. <https://doi.org/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. <https://doi.org/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. <https://doi.org/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. <https://doi.org/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. *Journal of Endocrinology*, 237, 193–205. <https://doi.org/10.1530/JOE-17-0642>
- Repouskou A, Panagiotidou E, Panagopoulou L, Bisting PL, Tuck AR, Sjodin MOD, Lindberg J, Bozas E, Ruegg J, Gennings C, Bornehag CG, Damdimopoulou P, Stamatakis A and Kitraki E, 2019. Gestational exposure to an epidemiologically defined mixture of phthalates leads to gonadal dysfunction in mouse offspring of both sexes. *Scientific Reports*, 9. <https://doi.org/10.1038/s41598-019-42377-6>
- Sandy EH, Yao J, Zheng SX, Gogra AB, Chen HL, Zheng H, Yormah TBR, Zhang X, Zaray G, Ceccanti B and Choi MMF, 2010. A comparative cytotoxicity study of isomeric alkylphthalates to metabolically variant bacteria. *Journal of Hazardous Materials*, 182, 631–639. <https://doi.org/10.1016/j.jhazmat.2010.06.079>
- Santangeli S, Maradonna F, Zanardini M, Notarstefano V, Gioacchini G, Forner-Piquer I, Habibi H and Carnevali O, 2017. Effects of diisononyl phthalate on Danio rerio reproduction. *Environmental Pollution*, 231, 1051–1062. <https://doi.org/10.1016/j.envpol.2017.08.060>
- Sarachana T, Xu MY, Wu RC and Hu VW, 2011. Sex Hormones in Autism: androgens and Estrogens Differentially and Reciprocally Regulate RORA, a Novel Candidate Gene for Autism. *Plos One*, 6. <https://doi.org/10.1371/journal.pone.0017116>
- Schwarz E, Guest PC, Rahmoune H, Wang L, Levin Y, Ingudomnukul E, Ruta L, Kent L, Spain M, Baron-Cohen S and Bahn S, 2011. Sex-specific serum biomarker patterns in adults with Asperger's syndrome. *Mol Psychiatry*, 16, 1213–1220. <https://doi.org/10.1038/mp.2010.102>
- Scott HM, Mason JI and Sharpe RM, 2009. Steroidogenesis in the Fetal Testis and Its Susceptibility to Disruption by Exogenous Compounds. *Endocrine Reviews*, 30, 883–925. <https://doi.org/10.1210/er.2009-0016>
- Sheikh IA, Turki RF, Abuzenadah AM, Damanhour GA and Beg MA, 2016. Endocrine Disruption: computational Perspectives on Human Sex Hormone-Binding Globulin and Phthalate Plasticizers. *Plos One*, 11:e0151444. <https://doi.org/10.1371/journal.pone.0151444>
- Somboonporn W and Davis SR, 2004. Testosterone Effects on the Breast: implications for Testosterone Therapy for Women. *Endocrine Reviews*, 25, 374–388. <https://doi.org/10.1210/er.2003-0016>
- Stroustrup A, Bragg JB, Andra SS, Curtin PC, Spear EA, Sison DB, Just AC, Arora M and Gennings C, 2018. Neonatal intensive care unit phthalate exposure and preterm infant neurobehavioral performance. *Plos One*, 13. <https://doi.org/10.1371/journal.pone.0193835>
- Takahashi K, Hosoya T, Onoe K, Takashima T, Tanaka M, Ishii A, Nakatomi Y, Tazawa S, Takahashi K, Doi H, Wada Y and Watanabe Y, 2018. Association between aromatase in human brains and personality traits. *Scientific Reports*, 8. <https://doi.org/10.1038/s41598-018-35065-4>
- van den Driesche S, Walker M, McKinnell C, Scott HM, Eddie SL, Mitchell RT, Seckl JR, Drake AJ, Smith LB, Anderson RA and Sharpe RM, 2012. Proposed Role for COUP-TFII in Regulating Fetal Leydig Cell Steroidogenesis, Perturbation of Which Leads to Masculinization Disorders in Rodents. *Plos One*, 7, e37064. <https://doi.org/10.1371/journal.pone.0037064>
- Wang R, Xu XH, Weng HF, Yan SY and Sun YY, 2016a. Effects of early pubertal exposure to di-(2-ethylhexyl) phthalate on social behavior of mice. *Hormones and Behavior*, 80, 117–124. <https://doi.org/10.1016/j.yhbeh.2016.01.012>
- Wang R, Xu XH and Zhu QJ, 2016b. Pubertal exposure to di-(2-ethylhexyl) phthalate influences social behavior and dopamine receptor D2 of adult female mice. *Chemosphere*, 144, 771–1779. <https://doi.org/10.1016/j.chemosphere.2015.10.062>
- Wang X, Wang L, Zhang JF, Yin WJ, Hou J, Zhang YJ, Hu C, Wang GY, Zhang R, Tao Y and Yuan J, 2018. Dose-response relationships between urinary phthalate metabolites and serum thyroid hormones among waste plastic recycling workers in China. *Environmental Research*, 165, 63–70. <https://doi.org/10.1016/j.envres.2018.04.004>
- Zhao Y, Ao H, Chen L, Sottas CM, Ge RS, Li LX and Zhang YH, 2012. Mono-(2-ethylhexyl) phthalate affects the steroidogenesis in rat Leydig cells through provoking ROS perturbation. *Toxicology in Vitro*, 26, 950–955. <https://doi.org/10.1016/j.tiv.2012.04.003>
- Zhao J, Ren S, Liu CY, Huo L, Liu Z and Zhai LL, 2018. Di-(2-Ethylhexyl) phthalate increases obesity-induced damage to the male reproductive system in mice. *Oxidative Medicine and Cellular Longevity*. <https://doi.org/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. <https://doi.org/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
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