Regulatory Toxicology and Pharmacology 81 (2016) 397–406

Contents lists available at ScienceDirect

Regulatory Toxicology and Pharmacology

journal homepage: www.elsevier.com/locate/yrtph

Assessment of reproductive and developmental effects of DINP, DnHP and DCHP using quantitative weight of evidence

Wolfgang Dekant [a, *], James Bridges [b]

[a] Department of Toxicology, University of Würzburg, 97078 Würzburg, Germany
[b] University of Surrey, Guildford, UK

ARTICLE INFO

Article history:
Received 20 May 2016
Received in revised form
22 September 2016
Accepted 27 September 2016
Available online 28 September 2016

Keywords:
Weight of evidence
Classification and labeling
Reproductive and developmental toxicity
Quantitative assessments

ABSTRACT

Quantitative weight of evidence (QWoE) methodology utilizes detailed scoring sheets to assess the quality/reliability of each publication on toxicity of a chemical and gives numerical scores for quality and observed toxicity. This QWoE-methodology was applied to the reproductive toxicity data on diisononylphthalate (DINP), di-n-hexylphthalate (DnHP), and dicyclohexylphthalate (DCHP) to determine if the scientific evidence for adverse effects meets the requirements for classification as reproductive toxicants. The scores for DINP were compared to those when applying the methodology DCHP and DnHP that have harmonized classifications. Based on the quality/reliability scores, application of the QWoE shows that the three databases are of similar quality; but effect scores differ widely. Application of QWoE to DINP studies resulted in an overall score well below the benchmark required to trigger classification. For DCHP, the QWoE also results in low scores. The high scores from the application of the QWoE methodology to the toxicological data for DnHP represent clear evidence for adverse effects and justify a classification of DnHP as category 1B for both development and fertility. The conclusions on classification based on the QWoE are well supported using a narrative assessment of consistency and biological plausibility.

1. Introduction

The process of hazard assessment and risk characterization should include a science-based evaluation of all of the available data on the investigation of the toxicity of a specific chemical (Beronius et al., 2014; Lutter et al., 2015; Rhomberg, 2015; Schreider et al., 2010; US-EPA, 2005). Traditionally, hazard assessments and risk characterization have relied on scientific judgment with a narrative assessment and have often only included results from "key studies". Consequently, the process has been criticized for lack of objectivity and transparency (see for example, Myers et al., 2009). However, conclusions based on the overall toxicity database often require integration of several lines of evidence (different types of studies) with different research objectives, applied methodologies and study quality. The available database may include peer-reviewed publications often addressing selected endpoints with potential relevance to toxicity, but also reports on the results of targeted toxicity testing following specific protocols required by legislation. In addition, the relevance of effects reported in scientific publications may be controversial. Therefore, narrative assessments have a number of weaknesses. To improve their quality, weight of evidence (WoE) approaches are increasingly mandated in chemical regulations (Agerstrand et al., 2014; ECHA, 2015; Weed, 2005). However, detailed guidance to perform WoE assessments is lacking and quantitative aspects only received limited considerations (Rhomberg, 2015; Van Der Kraak et al., 2014).

A recently developed quantitative weight of evidence (QWoE) approach to assess toxicity data for chemicals is designed to assist with classification and labeling (C&L) regarding reproductive toxicity endpoints (Dekant and Bridges, 2016). This QWoE applies predefined scoring criteria for relevant aspects of quality/reliability of a study for all reported effects to provide a fully transparent assessment. The scores representing strength of evidence for adverse effects are then compared to benchmark scores that are anchored to adverse biological endpoints and serve as the basic requirements for classification.

This QWoE was used to assess a need for C&L regarding findings from reproductive toxicity studies of three phthalates, diisononlyphthalate (DINP), dicyclohexylphthalate (DCHP), and di-n-hexylphthalate (DnHP). Phthalates are widely used as plasticizers.

* Corresponding author. Department of Toxicology, University of Würzburg, Versbacherstrasse 9, 97078 Würzburg, Germany.
E-mail address: dekant@toxi.uni-wuerzburg.de (W. Dekant).

http://dx.doi.org/10.1016/j.yrtph.2016.09.032

Glossary of terms

Weight of evidence method (WoE)	The identification and objective analysis (using predefined, scientifically justified criteria) of all potentially relevant studies, for their quality and in testing a hypothesis (problem formulation)
Quantitative weight of evidence (QWoE)	The identification, objective analysis and numerical scoring (using predefined scientifically justified criteria) of all potentially relevant studies, for both their quality and relevance in testing a hypothesis (problem formulation)
The hypothesis	Generally, takes the form of "does chemical of interest X cause adverse effects Y under conditions Z". Conditions may include exposure levels and duration, species of interest, adverse effects are defined as by WHO/IPCS
Endpoints	The measured and modelled findings used to identify and characterise adverse effects Y
Quality	The reliance that can be placed on the findings of each study for the purpose of critically testing the hypothesis
Relevance	The utility of the findings of each study on adverse endpoints for the purpose of critically testing the hypothesis
Lines of evidence	The different types of investigation used to critically test the hypothesis (e.g. observations in man, targeted toxicity testing in animals, in vitro experiments determining molecular endpoints, and in silico predictions of toxicity based on read-across or quantitative structure activity relationships)
Weighting of endpoints	A multiplier that is applied to the relevance/effect scores to reflect the relative importance of different types of endpoint and/or different lines of evidence in support of the hypothesis
Strength of evidence	This score is derived by multiplying the final relevance/effects scores by the quality/reliability score for a particular study
Overall weight of evidence	This is a summation of the findings from all suitable studies. It may be presented graphically as a plot of relevance/effects against quality scores or as an average numerical value with ranges

Di(2-ethylhexyl)phthalate (DEHP) was a plasticizer of choice, but has been replaced by higher molecular weight phthalates such as DINP in many applications due to concerns regarding possible adverse effects of DEHP-exposures in humans. A concern for phthalates in general, with a focus regarding adverse reproductive and developmental effects, has been raised based on results from animal toxicity studies with certain low molecular weight phthalates. In rats, application of high doses of di-n-butylphthalate (DBP) and DEHP during specific phases of pregnancy induce reproductive toxicity in male offspring (EFSA, 2005a; EFSA, 2005b; EFSA, 2005c; EFSA, 2005d; EFSA, 2005e). The effects include malformations of the epididymis, vas deferens, seminal vesicles, prostate, external genitalia (hypospadias) and cryptorchidism, as well as retention of nipples/areola (sexually dimorphic structure in rodents) and demasculinization of the perineum resulting in a reduction in anogenital distance (AGD). This pattern is sometimes termed the "phthalate syndrome" and is speculated as similar to a human disease termed 'testicular dysgenesis syndrome' (TDS). TDS in humans is hypothesized (Juul et al., 2014; Main et al., 2010; Sharpe and Skakkebaek, 2008) to account for many common disorders of newborn (such as cryptorchidism and hypospadias) and young adult males (such as low sperm count and testicular germ cell cancers) but the mode of action and underpinnings of TDS are unclear,. This concern on phthalates has raised a discussion on safety regarding many applications and resulted in national strategies regarding replacement of phthalates in commerce. However, the reproductive toxicity of low to medium molecular weight phthalates is different with clear effects observed for DEHP, DBP, and di(isobutyl)phthalate (DiBP) in one- and/or multigeneration studies (EFSA, 2005a; EFSA, 2005b; EFSA, 2005c; EFSA, 2005d; EFSA, 2005e; EU-RAR, 2008) whereas high molecular weight phthalates such as diisononylphthalate (DINP) and diisodecylphthalate (DIDP) did not induce such effects and reproductive toxicity is not considered a concern with dimethyl (DMP) and diethylphthalate (DEP) (Anonymous, 1997; Field et al., 1993; Gray et al., 2000; Hushka et al., 2001; SCCP, 2007; Waterman et al., 1999, 2000).

The purpose of the application of a QWOE to the toxicity database on DINP, DCHP, and DnHP was to assess the robustness of the QWOE-methodology and the relevance of reported effects in the scientific literature in a transparent, consistent and scientifically justified way, using predetermined scores for quality and relevance/effects. DnHP and DCHP have harmonized classifications according to the CLP regulation as category 1B reproductive toxicants (DnHP for both development and fertility; DCHP only for development), while, according to the European Risk Assessment report (EU-RAR, 2003), a classification of DINP was not mandated at the time of the preparation of the EU-RAR. Since completion of the EU RAR, little new information on the effects of DINP on reproductive endpoints has been generated and is integrated here.

2. Methods

In the first step, potentially useful publications for assessment purposes on the animal toxicology of DINP, DCHP, and DnHP were searched with a cut-off date of July 31, 2015. To capture all publications and minimize search-bias, the literature search included PubMed, TOXLINE, Chemical Abstracts, and SciFinder with the following search terms:

- CAS # 84-61-7 (DCHP), CAS # 84-75-3 (DnHP), CAS # 28553-12-0, and CAS # 68515-48-0 (both DINPs) in "ToxLine",
- CAS # 84-61-7, CAS # 84-75-3, CAS # 28553-12-0, and CAS # 68515-48-0 and "toxicity" in "Chemical abstracts"
- CAS # 84-61-7, CAS # 84-75-3, CAS # 28553-12-0, and CAS # 68515-48-0 and "toxicity" in "PubMed"

- CAS # 84-61-7, CAS # 84-75-3, CAS # 28553-12-0, and CAS # 68515-48-0 and “toxicity” in “SciFinder”

All publications identified by these searches are listed in Annex 1 provided as supplementary material. The results of the literature search were compared to reports and publications listed in the EU risk assessment report on DINP (EU-RAR, 2003). In addition, original study reports of guideline studies on DINP were available. Current understanding of the biology and modes of action involved in reproductive toxicology was obtained from appropriate textbooks and reviews (Gupta, 2011; Klaassen, 2013). Abstracts of all publications identified in the literature search were read. To narrow the focus of the evaluations, all publications that did not address results from animal toxicity studies on the phthalates of interest were not further considered. All animal toxicity studies were then screened to identify if they addressed endpoints relevant to reproductive toxicity in their design and effects assessment (see below for details). Only publications that fulfilled this criterion were evaluated further. The detailed methodology including detailed scoring criteria for study quality and relevance/effects can be found in Dekant and Bridges (2016). In brief, the developed QWoE methodology utilizes numerical scoring to assess reliability of a publication and the toxicological relevance of reported effects. Scores are given for fourteen quality aspects, best practice receives the highest score. The relevance/effects scores (0 to four) are adjusted to the key elements of the toxic response for the endpoint and include weighting factors for effects on different levels of biological organization. The relevance/effects scores are then assessed against the criteria dose-response, magnitude and persistence of effects, consistency of observations with the hypothesis, and relation of effects to human disease. The quality/reliability scores and the relevance/effect scores are then multiplied to give a numerical strength of evidence for adverse effects. This total score is then used to assign the chemical to the different classes employed in classification (Dekant and Bridges, 2016).

3. Results

3.1. Application of the QWoE-methodology to the database

The QWoE methodology was used to score all available publications for quality/reliability of results and for relevance/strength of effects. Scoring sheets were defined as described (Dekant and Bridges, 2016). The quantitative scores were derived for fourteen aspects relevant to quality, with the best practice for each specific aspect receiving the highest score of four. A relevance of effect score was given for each observation in a study. This score consisted of two parts, a nature of effect score that places different weights on the various types of observations based on their biological proximity to an apical outcome and a strength of evidence score consisting of five criteria aimed at quantifying the toxicological significance of the observation (Dekant and Bridges, 2016). The quality scores and the relevance/strength of effects scores were multiplied to give a numerical strength of evidence for adverse effects on reproduction and on fertility. The total score for each publication was brought together as a conclusion of the total database for each chemical of interest and used to assign each chemical to the different classes employed in classification (ECHA, 2015). The results are presented in a table submitted as Annex 2 in the supplementary material.

Because the details of the developed score sheets were novel, and there was no prior experience of their use, it was crucial to conduct independent assessments for a range of the publications to compare the scoring results. To check consistency in the scoring, 18 publications (ten for DINP, four for DCHP, and four for DnHP) were independently scored by both authors. A good general agreement for both quality and relevance/strength scores was documented (for a tabulated overview of the initial scoring, see Annex 2). Quality and relevance/strength of effect scores only differed by app. 20% and neither assessor had a general tendency to assign higher or lower scores than the other. This demonstrates that criteria specified in the scoring sheets are interpreted identically by both assessors. In those few studies where there were substantial variations, the scoring guidance sheets were re-reviewed to identify the cause of discrepancies. Once this was completed, double scoring of all available papers on DINP, DCHP and DnHP was not considered necessary and only one assessor (WD) scored all studies. The few studies not scored by JB are indicated by NS in Annex 2.

3.2. Conclusions on the overall strength of evidence to support a Cat 1B or 2 for reproductive and developmental toxicity according to CLP for DINP, DCHP, and DnHP

To develop conclusions regarding support for classification and labeling, general developmental/reproductive effects and effects on fertility were treated separately because the regulation on classification separates these endpoints. The QWoE presented here is quantitative and is developed specifically for application to hazard classification mandated by EC regulation 1272/2008 (EC-Regulation, 2008) on classification, labeling and packaging of substances and mixtures (i.e the CLP Regulation or CLP).

3.2.1. Studies addressing developmental effects

General developmental toxicity was addressed by studies investigating effects of DINP, DCHP and DnHP in offspring including those with a follow-up of offspring until adulthood after in utero exposure. The database thus included ten studies on DINP, five studies on DCHP, and four studies on DnHP (Table 1). For studies on developmental toxicity of DINP, the quality scores ranged from 2.0 to 3.79 (Table 1). For both DCHP and DnHP, average quality scores were in the same range as for DINP, the databases on DCHP and DnHP are much more limited with only few high quality studies available (Table 1). It may be concluded that the available databases on DINP, DCHP, and DnHP have an acceptable average quality but individual study quality varies widely. High quality scores were assigned to studies where the animal model was appropriate, adequate numbers of animals/dose group were used with suitable dose spacing and timing, concurrent controls were examined, information was provided on historic controls, pathological and gross assessments were properly performed (e.g. demonstrated expertise and blinded readings/measurements), and the statistical evaluation was appropriate (Clewell et al., 2013a, 2013b; Exxon, 1994, 1996a, 1996b; Waterman et al., 1999; Waterman et al., 2000).

Lower scores regarding quality were usually due to combinations of the following issues:

- No access to raw data/only limited data provided
- major weaknesses in study design or execution, such as low numbers of animals/dose group and/or inappropriate timing of dosing or sampling
- inadequate characterization of the test chemicals and their concentration and stability in the application medium
- limited or absent descriptions of quality control and measurement methodology.
- absence of measures taken to avoid contamination from potentially interfering chemicals
- inadequate assessment of maternal toxicity (only maternal weight gain was determined in many studies, despite information that hepatotoxicity may occur with the studied phthalates

Table 1
Overview on studies assessing potential adverse effects on development in animals for DINP, DCHP and DnHP used for final scoring regarding level of evidence of adverse effects (NS = not scored). The table includes scoring of all publications identified in the literature search addressing developmental endpoints.

Publication	Quality JB	Quality WD	Effects JB	Effects WD	Overall score JB	Overall score WD
Summary of scores for DINP papers and reports						
(Boberg et al., 2011)	2.29	2.14	34.00	40.00	77.86	85.60
(Clewell et al., 2011a; Clewell et al., 2013b)	3.54	3.79	37.50	33.00	132.75	125.07
(Clewell et al., 2011b; Clewell et al., 2013a)	3.28	3.57	21.00	29.00	68.88	103.53
(Masutomi et al., 2003)	2.92	2.15	49.00	64.00	143.08	137.60
(Exxon, 1996a)	NS	3.58	NS	16.00	NS	57.28
(Exxon, 1996b)	3.36	3.75	24.00	16.00	80.64	60.00
(Exxon, 1994)	3.09	3.45	10.5	6.00	32.45	20.70
(Hellwig et al., 1997)	NS	3.45	NS	0.00	NS	0.00
(Li et al., 2015)	2.08	2.07	26.00	71.00	54.08	146.97
(Gray et al., 2000)	2.60	2.57	48.00	66.00	124.80	169.62
Average overall score	**2.90**	**3.05**	**31.25**	**34.10**	**89.32**	**90.64**
Summary of scores for DCHP papers and reports						
(Saillenfait et al., 2009b)	3.08	3.38	51.00	24.00	157.08	81.12
(Aydogan Ahbab and Barlas, 2013)	2.08	1.79	72.50	115.50	150.80	206.75
(Aydogan Ahbab and Barlas, 2015)	2.15	1.86	72.00	84.50	154.80	157.17
(Hoshino et al., 2005)	2.86	3.21	96.00	109.50	274.56	351.50
(Yamasaki et al., 2009)	NS	2.77	NS	90.50	NS	250.69
Average overall score	**2.54**	**2.60**	**72.88**	**84.80**	**184.31**	**209.44**
Summary of scores for DnHP papers and reports						
(Saillenfait et al., 2009b)	3.09	3.23	186.00	193.00	574.74	623.39
(Saillenfait et al., 2009a)	3.08	3.38	139.00	160.50	428.12	542.49
(Aydogan Ahbab and Barlas, 2013)	2.08	1.79	72.50	108.50	150.80	194.22
(Aydogan Ahbab and Barlas, 2015)	2.15	1.86	72.00	90.00	154.80	167.40
Average overall score	**2.60**	**2.57**	**117.38**	**138.00**	**327.12**	**381.87**

in maternal animals under the dosing conditions (ECHA/RAC, 2013)

Relevance/effects scores were developed based on the effects reported in the studies and include weighting factors for effects on different levels of biological organization. In the QWoE-approach, weighted endpoints were checked against the criteria dose-response, magnitude and persistence of effects, consistency of observations with the hypothesis, and relation of effects to human disease (Dekant and Bridges, 2016). Detailed results of the QWoE-evaluation are given in the scoring sheets in Annex 3. Relevance/effect scores regarding developmental effects for DINP ranged from zero to 71, relevance/effects scores for DCHP ranged from 24 to 115.5, and from 72 to 193 for DnHP (Table 1).

3.2.2. Scoring for effects on fertility

Regarding fertility, the following types of studies were considered:

- exposure of both male and female animals over at least two weeks before mating and during the mating period.
- consequences of in utero exposures in offspring after reaching adulthood.
- "classical" repeated dose studies assessing reproductive organ function and histopathology.

The database evaluated included five studies on DINP, three studies on DCHP, and three studies on DnHP. Thus, the number of studies assessing fertility endpoints for the three phthalates of interest was more limited compared to those assessing developmental effects. The results of the scoring for quality and relevance/strengths of effects of all studies addressing fertility endpoints are summarized in Table 2. Individual scoring sheets are detailed in Annex 3. As with the studies on developmental effects, based on the scoring for quality, the available databases on DINP, DCHP, and DnHP are generally of acceptable quality. Again, the quality of the individual studies varies widely, In common with the developmental effects, relevance/strength of effects scores were highly different for the three phthalates of interest with DINP receiving low effects scores while DnHP had high effect scores due to clear adverse effects.

A plot of quality scores versus relevance/effects scores gives a graphical overview on the results of the QWoE and permits an assessment of distribution of scores and potential uncertainties (Figs. 1 and 2). When comparing the score plots for developmental effects (top chart, Fig. 1), it is evident that most high quality studies do not provide any evidence for adverse effects of DINP on development since the highest quality studies consistently received relevance/effects scores well below 50. The database remains consistent when integrating all studies with an acceptable quality (score above 2). The more limited database on developmental effects for DCHP (center, Fig. 1) shows an acceptable quality for most studies, but a larger spread of relevance/effects scores with several high quality studies reaching scores of 100 and above. For DnHP (lower chart, Fig. 1), quality is also acceptable for most studies with two high quality studies receiving relevance/effect scores of >150.

Plots of the scoring regarding fertility also consistently show no evidence for adverse effects from the studies on DINP with the highest quality studies receiving very low scores for relevance/effects (Fig. 2, top chart). The databases on DCHP and DnHP regarding the endpoint fertility are more limited (Fig. 2, center and bottom chart) and show a larger spread of quality scores due to the presence of studies with lower reliability with DCHP receiving top scores of >100 and one high quality study on DnHP with a score of >200 indicating strong evidence for adverse effects.

4. Discussion

The criteria in the CLP Regulation (EC Regulation 1272/2008 1272/2008 part 3.7) were used as a basis for conclusions based on the QWoE regarding C&L of DINP, DCHP, and DnHP (Dekant and Bridges submitted manuscript). According to the classification criteria, classification has to be based on the induction of adverse effects in appropriately performed animal toxicity studies. However, according to the guidance, even in the presence of adverse effects, classification is not appropriate if:

Table 2
Overview on studies assessing potential adverse effects on fertility in animals for DINP, DCHP and DnHP used for final scoring regarding level of evidence of adverse effects (NS = not scored). The table includes scoring of all publications identified in the literature search addressing fertility.

Publication	Quality JB	Quality WD	Effects JB	Effects WD	Overall score JB	Overall score WD
Summary of scores for DINP papers and reports						
(Boberg et al., 2011)	2.29	2.14	15.00	6.00	34.35	12.84
(Masutomi et al., 2003)	2.92	2.15	52.00	69.50	151.84	149.43
(Kwack et al., 2009)	2.18	2.00	32.50	10.50	70.85	21.00
(Exxon, 1996a)	NS	3.58	NS	0.00	NS	0.00
(Exxon, 1996b)	3.36	3.75	16.50	0.00	55.44	0.00
Average overall score	**2.69**	**2.72**	**29.00**	**17.20**	**78.12**	**36.65**
Summary of scores for DCHP papers and reports						
(Aydogan Ahbab and Barlas, 2013)	2.08	1.79	125.00	96.00	260.00	171.84
(Hoshino et al., 2005)	2.86	3.21	92.50	60.50	264.55	194.21
(Yamasaki et al., 2009)	NS	2.77	NS	67.50	NS	186.98
Average overall score	**2.47**	**2.59**	**108.75**	**74.67**	**262.28**	**184.34**
Summary of scores for DnHP papers and reports						
(Aydogan Ahbab and Barlas, 2013)	2.08	1.79	125.00	87.50	260.00	156.63
(Lamb et al., 1987)	NS	3.00	NS	208.00	NS	624.00
(Foster et al., 1980)	NS	2.00	NS	64.50	NS	129.00
Average overall score	**2.08**	**2.26**	**125.00**	**120.00**	**260.00**	**303.21**

i) The adverse effect is produced solely as a non-specific secondary consequence of other toxic effects (e.g. due to maternal toxicity).
ii) There is an adequate number of studies but insufficient strength of evidence when considered together (3.7.2.3.1).
iii) Differences in toxicokinetics are so marked that it is certain that hazardous effects seen in the animal model will not be seen in man (3.7.2.3.2, 3.72.5.8).
iv) Mode of action differences are so marked that it is certain that hazardous effects seen in the animal model will not be seen in man (3.7.2.3.3).
v) The effects observed are considered to be of low or minimum significance (e.g. small changes in semen parameters, small changes in the incidence of spontaneous defects in the fetus (3.7.2.3.3).
vi) The study uses exposure levels above the limit dose (however, there is no clear definition but 1000 mg/kg bw/day is indicated as a benchmark in 3.7.2.5.7 of the guidelines).

WoE-approaches are included in the CLP Regulation (3.7.2.3.1) which state that "the weight given to the available evidence will be influenced by factors such as the quality of the studies, consistency of results, nature and severity of effects, the presence of maternal toxicity in experimental animal studies, level of statistical significance for inter-group differences, number of endpoints affected, relevance of route of administration to humans and freedom from bias. Both positive and negative results are assembled together into a weight of evidence determination." These essential issues are addressed by the QWoE methodology. Level of evidence for adverse effects is obtained by the scoring for relevance/strength of effects and addresses points i, ii, v, and vi outlined above.

Regarding conclusions on C&L, the overall strength of evidence obtained by a multiplication of the quality and relevance/strength scores is applied with numerical limits defined based on level of evidence (Dekant and Bridges, 2016). For chemicals where all studies have an overall weight of evidence score below 245, there is insufficient evidence of the induction of adverse effects and therefore no scientific basis for classification as a reproductive toxicant. Overall scores between 246 and 350 indicate limited evidence for adverse effects thus supporting a classification as a category 2 reproductive toxicant. Overall scores above 351 confirm strong evidence for induction of adverse effects and a classification as a category 1B reproductive toxicant (Dekant and Bridges, 2016).

One issue with deriving overall weight of evidence scores is the large range of endpoints assessed by the individual studies in a larger database. For the three phthalates assessed here, some studies only assessed biochemical parameters or observed minor changes with an unknown relation to adverse effects while others performed a comprehensive assessment of potential adverse effects. Scores for these studies are contained in Annex 2. Studies limited to biochemical parameters receive low relevance/effects scores. Inevitably, inclusion of studies with a limited scope that results in a low relevance/effects score will bias the summary results towards lower scores. For example, if the studies addressing only biochemical effects were to be included in the summary scores, this decreases the overall weight of evidence score from 90.64 (Table 1) to 68.24 (for WD who scored all studies), from 209.44 (WD) to 176.54 for DCHP, and from 381.87 (WD) to 273.38 for DnHP. It was therefore decided to exclude studies with very limited relevance/effects scores (Adamsson et al., 2009; Borch et al., 2004; Furr et al., 2014; Kwack et al., 2009) from calculating an overall weight of evidence score.

For DINP, all overall weight of evidence scores for development and the overall weight of evidence scores for all individual studies (Fig. 3) remain well below the threshold for "limited evidence to support induction of adverse effects" (Dekant and Bridges, 2016). Scores for adverse developmental effects of DCHP and DnHP both for individual studies and the overall weight of evidence scores (Fig. 3) are much higher. For DCHP, the overall score for effects on development remains below the threshold for "limited evidence", but scores for one study by one of the assessors reaches the threshold for "strong evidence" with two additional scores supporting "limited evidence" (Fig. 3, top chart). DnHP reached an overall score above the threshold to conclude on "good evidence" for adverse developmental effects (Fig. 3, top chart). For fertility endpoints (Fig. 3, lower chart), all scores for DINP also provide no evidence of adverse effects. Overall, the database for DCHP also does not provide evidence for an adverse effect on fertility, although the strength of evidence scores for one study just passes the threshold for "limited evidence". The scoring of effects of DnHP on fertility on average (Fig. 3, lower chart) only supports "limited evidence for an adverse effect". However, one of the high quality publications on DnHP indicates "good/strong evidence of an adverse effect". Effects observed in this study are consistent with observations made with other low to medium molecular weight phthalates, are biologically plausible and consistent with the hypothesis. Therefore, the QWoE supports a classification of DnHP into category 1b for fertility.

In summary, the QWoE evaluation showed clear differences in the overall scores between DINP and the benchmark chemicals

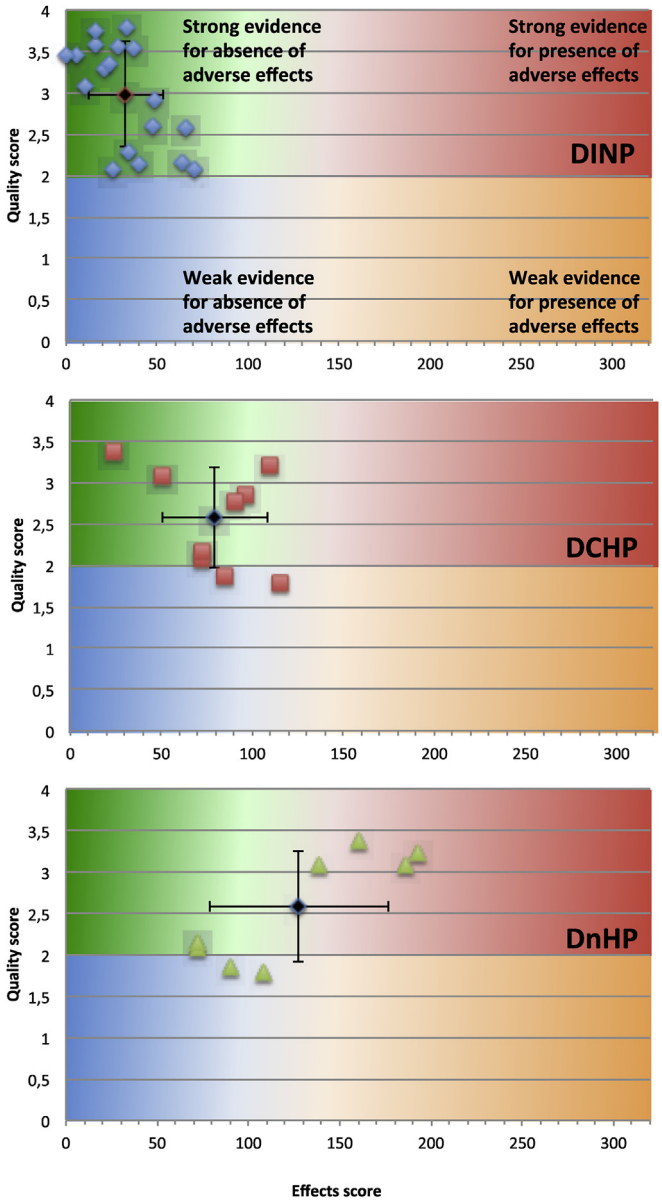

Fig. 1. Plot of quality/reliability scores versus relevance/effect scores for the outcome of the QWoE assessment of all studies addressing developmental endpoints of DINP, DCHP, and DnHP. Figure includes means ± SD (◆) of scores for each of the phthalates.

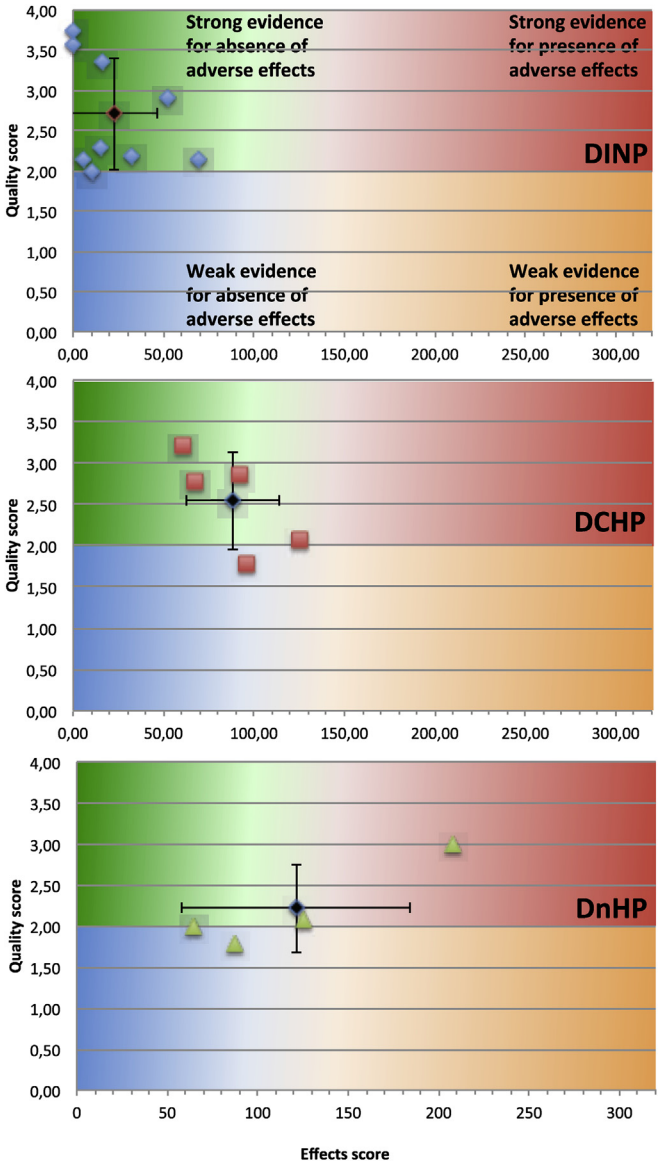

Fig. 2. Plot of quality/reliability scores versus relevance/effect scores for the outcome of the QWoE assessment of all studies addressing fertility endpoints of DINP, DCHP, and DnHP. Figure includes means ± SD (◆) of scores for each of the phthalates.

DCHP and DnHP (Fig. 3). Based on the scores and the overall weight of evidence, there is no biologically justified support for an assignment of a CLP reproductive classification category for DINP. Using a conservative approach, the QWoE supports a classification of DCHP as a category 1B reproductive toxicant regarding developmental effects due to scores above the threshold for one good quality study and no support for classification on fertility. However, based on the overall weight of evidence scoring, category 2 for development is considered more appropriate. The outcome of the QWoE for DnHP well supports the existing classification for DnHP as category 1b reproductive toxicant regarding both development and fertility.

4.1. Comparison of the results of the QWoE with narrative assessments for DINP

For a further validation of the QWoE approach, a comparison of the outcome of the QWoE with a detailed narrative approach regarding effects reported in the databases on DINP, DCHP, and DnHP is detailed below.

4.1.1. Developmental effects

The available DINP animal toxicity studies show that a "classical" narrative approach well supports the conclusions from the QWoE. DINP (Waterman et al., 1999, 2000) does not cause any of the permanent effects that have been observed with short and medium chain phthalates (EFSA, 2005a). For example, malformations and permanent histopathological changes of the male reproductive tract were not observed. Effects on reproductive organs in newer studies (Boberg et al., 2011; Clewell et al., 2013a, 2013b) were only transient, did not persist to adulthood, and occurred at doses levels where a number of other studies with DINP have demonstrated maternal toxicity. Moreover, some of the effects such as AGD and nipple retention (Boberg et al., 2011) are only significant at the $p < 0.05$ level and reported the dose-dependency

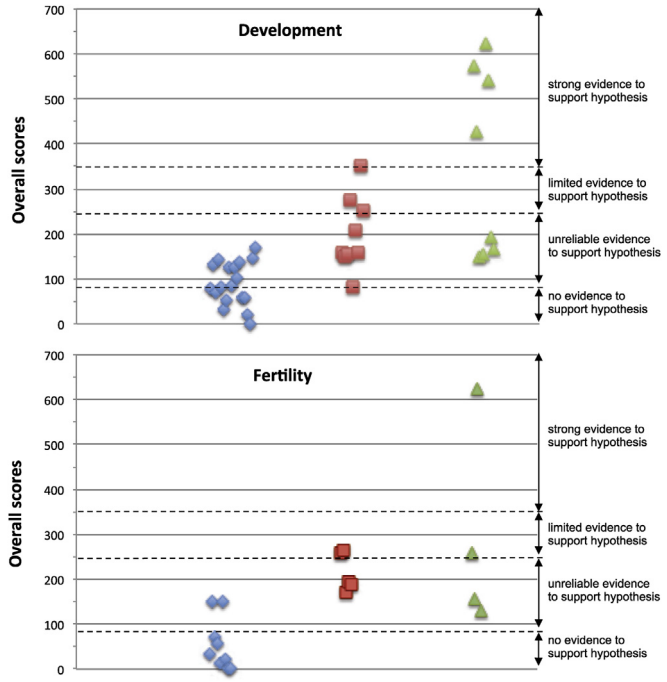

Fig. 3. Distribution of strength of evidence score for DINP, DCHP, and DnHP for developmental (top chart) and fertility endpoints (bottom chart) and relation to thresholds regarding strength of evidence and basis for classification. Limited evidence to support the hypothesis of induction of adverse effects requires a minimum score of 246 and strong evidence a minimum score of 350. Limited evidence supports a classification as a category 2 reproductive toxicant and strong evidence supports a classification as a category 1B reproductive toxicant.

is difficult to rationalise bearing in mind the saturation of DINP bioavailability at doses > 300 mg/kg bw (Clewell et al., 2013a). Therefore, it appears unlikely that the reported changes are related to an increase in foetal exposure to DINP or DINP-metabolites.

An increased incidence of "malformations" following DINP administration was only reported in one study with DINP (Gray et al., 2000), but statistical significance was only reached when the different types of effects were summed for statistical analysis. This is inappropriate. In addition, this study reported an increase in areolae at PND13 (22.4% in DINP-treated animals as compared to an incidence of 0% in controls), but a later study from the same laboratory reported an areolae incidence of 14% in controls (Ostby et al., 2001). Both studies were evaluated regarding consequences for classification (EU-RAR, 2003) and were not considered as providing information to warrant classification.

The toxicological significance of the reversible induction of multinucleated germ cells (MNGs) (Boberg et al., 2011; Clewell et al., 2011a, 2011b) DINP is unclear. Reduced testicular testosterone levels are unlikely to be involved in the etiology of MNG formation in rats (Scott et al., 2007). MNGs are also apparently not formed by proliferative events (Spade et al., 2015). In summary, these findings suggest that MNG formation is not androgen dependent and thus may not represent an endpoint that is consistent with the hypothesis that an interference with testosterone in rats is a key effect in the reproductive toxicity of some phthalates. Moreover, the increased incidences of MNGs apparently have no functional consequences (Waterman et al., 1999, 2000). Thus, they do not qualify as being adverse. Nonetheless, in the QWoE assessment developed, these observations were included in the weighted scores. The available DINP developmental toxicity studies also showed some increases in the incidences of minor skeletal variations at maternally toxic dose levels, but no abnormalities were reported. Likely, these variations are induced by maternal effects of the very high doses of DINP. The available developmental toxicity studies were evaluated in the EU-RAR in 2003 with the conclusion "*Regarding …. development, the effects observed in the available studies, do not justify classification according to the EU classification criteria*". The QWoE demonstrates that this conclusion remains valid.

4.1.2. Effects on fertility

Effects of DINP on fertility-related parameters were only reported in two studies (Boberg et al., 2011; Kwack et al., 2009). Kwack et al., 2009 reported significantly lowered sperm counts and motility of epididymal sperm after a four-week treatment of adult rats with a single dose level of 500 mg DINP/kg bw/day by daily gavage. The study suffers from the use of a single dose level and unclear reporting regarding use of concurrent controls. Moreover, the changes in sperm number only had a statistical significance level of <0.05 and general toxicity of DINP has been reported to occur at dose levels above ~200 mg/kg bw/day (ECHA/RAC, 2013). The limited information provided (Kwack et al., 2009) also indicates treatment-related general toxicity such as body weight reduction, increases in relative organ weights, effects on clinical chemistry and hematology parameters. Thus, the effects on sperm count and sperm motility may be secondary to systemic toxicity.

Boberg et al., 2011 assessed reproductive parameters in male rats exposed in utero to doses of up to 900 mg DINP/kg bw/day. Sperm analysis at PND90 showed an increase in sperm count only the highest dose and a statistically significant, but not dose-related decrease in sperm motility at the 600, 750, and 900 mg/kg bw/day. Testes weights and histology were unaffected at PND 90. The study has several limitations. Sperm analyses examined only 1–3 animals per litter and sperm motility in control was low and the changes noted in the DINP-treated groups remained within the range of historical controls (Jarfelt et al., 2005; Taxvig et al., 2007). According to the guidance (OECD, 2008), in general, 200 sperm should be analyzed and a minimum value of 70% motility is acceptable in controls. Further, dose-response and changes in testes histology are required to determine whether an effect is considered to be adverse. Testes histology is considered the most sensitive endpoint indicative of adverse effects of chemicals on male fertility (Mangelsdorf et al., 2003). Therefore, the information presented regarding sperm motility is inconclusive. The absence of effects on development and fertility in the two-generation reproductive and developmental toxicity study (Waterman et al., 2000) does not support the effects of DINP on fertility. While the design of these studies does not specifically address early life effects on male sexual development, it is noted that DEHP and other low molecular weight phthalates did cause persistent adverse effects on male reproductive performance in similar studies (EFSA, 2005a; EFSA, 2005c). While some of the available 90-day toxicity studies with DINP reported some effects on male reproductive organ weight, these effects occurred at much higher doses than those causing hepatotoxicity (the former ECB defined 276 mg/kg bw/day as NOAEL for effects of DINP on reproductive organ weight changes as compared to liver toxicity for which the NOAEL for liver effects of DINP has been identified by ECHA as 15 mg/kg bw/day). In summary regarding fertility, changes reported in the individual studies with DINP were inconsistent regarding incidences and periods of observed changes and no permanent changes were observed. Therefore, the transient effects seen (Boberg et al., 2011) do not qualify as "adverse" regarding "sexual function and fertility". It is therefore evident that there is no *"clear evidence of an adverse effect on sexual function… in the absence of other toxic effects"* and a classification of DINP into Cat. 1B or 2 according to CLP is neither supported by a narrative approach nor by the QWoE applied here.

This is consistent with the guidance in the CLP regulation (3.7.2.3.3) that states that "small changes in semen parameters or in the incidence of spontaneous defects in the fetus, small changes in the proportions of common fetal variants such as are observed in skeletal examinations, or in fetal weights, or small differences in postnatal developmental assessments" cannot be used to justify classification.

4.2. Comparison of the results of the QWoE with narrative assessments for DCHP

4.2.1. Developmental effects

For DCHP, only five studies on developmental toxicity endpoints are available. The average scores from the QWoE do not support a classification as a category 2 developmental toxicant. However, an individual study received a score of 274.56 by one assessor and of 351.5 by the other assessor, which is at the borderline for classification as category 1b (score of 351 and above). The low overall weight of evidence score of the database on DCHP regarding developmental effects is mainly due to the low scoring of two publications (Aydogan Ahbab and Barlas, 2013; Aydogan Ahbab and Barlas, 2015) that investigated relevant endpoints for DCHP and observed a number of changes that can be considered as adverse. However, the changes were observed at all dose levels of DCHP without an obvious dose-response despite applying a 25-fold spread between the tested doses (20 mg/kg/day to 500 mg/kg bw/day) and occurred at dose levels where neither of the two other more comprehensive studies (Hoshino et al., 2005; Saillenfait et al., 2009a) report DCHP-induced changes when compared to concurrent controls. However, when assessed individually, Hoshino et al., 2005 reports changes in reproductive organ weights and AGD/nipple development in male F1 and F2 pups. The 2nd high quality study for reproductive effects of DCHP (Saillenfait et al., 2009a) observed very limited effects for this endpoint due to the study focus on developmental changes and was not included in the overall weight of evidence scoring. In summary, the high scores of the high quality study support category1B when applying a conservative approach.

4.2.2. Effects on fertility

Of the three studies addressing fertility aspects of DCHP, only one can be considered of good quality. Only one QWoE score, 264.55 by one assessor, is above the threshold (246) suggesting "limited evidence" for an adverse effect on fertility. The overall score, however, remains well below "limited evidence". Hoshino et al., 2005 did not observe clear effects on fertility parameters in the two-generation study and only noted an approx. 20% decrease in homogenization resistant spermatids in the F1 parental generation, apparently without functional consequences. Yamasaki et al., 2009 describes the presence of decreased testicular germ cells in animals sacrificed at ten weeks of age in the high dose group of 500 mg DCHP/kg bw/day, but gives neither data on incidences nor a statistical evaluation (Yamasaki et al., 2009). Therefore, classification regarding fertility is not supported for DCHP by a narrative assessment.

4.3. Comparison of the results of the QWoE with narrative assessments for DnHP

4.3.1. Developmental effects

Only two studies were identified as high quality (Saillenfait et al., 2009a, 2009b). Both studies report a series of adverse effects that were persistent up to 14 weeks of age and thus support a classification regarding developmental effects as category 1b. The high scores of these two animal studies (>500) also indicate clear evidence of adverse effects in studies relevant for classification and labeling purposes. A 3rd study (Saillenfait et al., 2013) only assessed changes in testosterone and expression of proteins involved in steroid biosynthesis. The data in this study show a clear dose-dependent reduction of testicular testosterone (in contrast to data on DINP) and changes in steroid biosynthesis gene expression. These findings are consistent with the above conclusion; however, the study did not assess adverse changes. The overall score regarding developmental effects of DnHP again is reduced due to the two studies with a lower quality assignment (Aydogan Ahbab and Barlas, 2013; Aydogan Ahbab and Barlas, 2015). However, the clear effects on development well support a category 1B for development.

4.3.2. Effects on fertility

Work by the US NTP (published as Lamb et al., 1987) reports a marked effect of high doses of DnHP-administration on fertility in mice. In addition, Saillenfait et al. (2009b) also reports clear and permanent effects of DnHP-administration on reproductive organ histopathology. Moreover, short term administration of DnHP for 4 days has been shown to induce marked effects on testes weight in the absence of effects on body weight (Foster et al., 1980). Therefore, the narrative assessment of effects relevant to C&L for DnHP is consistent with the outcome of the QWoE assessment.

Acknowledgments

Preparation of this manuscript was supported in part through an honorarium to Drs. Dekant and Bridges from the ECPI. This review represents the individual professional views of the authors and not necessarily the views of the ECPI.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.yrtph.2016.09.032.

Transparency document

Transparency document related to this article can be found online at http://dx.doi.org/10.1016/j.yrtph.2016.09.032.

Conflicts of interest

Drs. Dekant and Bridges report personal fees from the European Council for Plasticizers and Intermediates (ECPI).

References

Adamsson, A., Salonen, V., Paranko, J., Toppari, J., 2009. Effects of maternal exposure to di-isononylphthalate (DINP) and 1,1-dichloro-2,2-bis(p-chlorophenyl) ethylene (p,p'-DDE) on steroidogenesis in the fetal rat testis and adrenal gland. Reprod. Toxicol. 28, 66–74.

Agerstrand, M., Edvardsson, L., Ruden, C., 2014. Bad reporting or bad Science? Systematic data evaluation as a means to improve the use of peer-reviewed studies in risk assessments of chemicals. Hum. Ecol. Risk Assess. 20, 1427–1445.

Anonymous, 1997. Reproductive toxicology: diethylphthalate. Reprod. Toxicol. 105, 245–246.

Aydogan Ahbab, M., Barlas, N., 2013. Developmental effects of prenatal di-n-hexyl phthalate and dicyclohexyl phthalate exposure on reproductive tract of male rats: postnatal outcomes. Food Chem. Toxicol. 51, 123–136.

Aydogan Ahbab, M., Barlas, N., 2015. Influence of in utero di-n-hexyl phthalate and dicyclohexyl phthalate on fetal testicular development in rats. Toxicol. Lett. 233, 125–137.

Beronius, A., Molander, L., Ruden, C., Hanberg, A., 2014. Facilitating the use of non-standard in vivo studies in health risk assessment of chemicals: a proposal to improve evaluation criteria and reporting. J. Appl. Toxicol. 34, 607–617.

Boberg, J., Christiansen, S., Axelstad, M., Kledal, T.S., Vinggaard, A.M., Dalgaard, M., Nellemann, C., Hass, U., 2011. Reproductive and behavioral effects of diisononyl phthalate (DINP) in perinatally exposed rats. Reprod. Toxicol. 31, 200–209.

Borch, J., Ladefoged, O., Hass, U., Vinggaard, A.M., 2004. Steroidogenesis in fetal male rats is reduced by DEHP and DINP, but endocrine effects of DEHP are not modulated by DEHA in fetal, prepubertal and adult male rats. Reprod. Toxicol. 18, 53–61.

Clewell, R., Andersen, M.E., Sochaski, M., 2011a. A Dose Response Study of the Effects on Male Rat Sexual Development after Administration of Diisononyl Phthalate to the Pregnant and Lactating Dam. Final Report N° 10003 (DiNP phase II study). Performing laboratory The Hamner Institutes for Health Sciences, Research Triangle Park, NC; Sponsor ExxonMobil Biomedical Sciences Inc, Annandale, NJ.

Clewell, R., Andersen, M.E., Sochaski, M., 2011b. Pharmacokinetics and Fetal Testes Effects after Diisononyl Phthalate Administration in Rat Gestation. Final Report N° 09016 (DiNP phase I study). Performing laboratory The Hamner Institutes for Health Sciences, Research Triangle Park, NC; Sponsor ExxonMobil Biomedical Sciences Inc, Annandale, NJ.

Clewell, R.A., Sochaski, M., Edwards, K., Creasy, D.M., Willson, G., Andersen, M.E., 2013a. Disposition of diisononyl phthalate and its effects on sexual development of the male fetus following repeated dosing in pregnant rats. Reprod. Toxicol. 35, 56–69.

Clewell, R.A., Thomas, A., Willson, G., Creasy, D.M., Andersen, M.E., 2013b. A dose response study to assess effects after dietary administration of diisononyl phthalate (DINP) in gestation and lactation on male rat sexual development. Reprod. Toxicol. 35, 70–80.

Dekant, W., Bridges, J., 2016. A quantitative weight of evidence methodology for the assessment of reproductive and developmental toxicity and its application for classification and labeling of chemicals. Regul. Toxicol. Pharmacol. http://dx.doi.org/10.1016/j.yrtph.2016.09.009.

EC-Regulation, 2008. Regulation (EC) No 1272/2008 of the European Parliament and of the Council of 16 December 2008 on Classification, Labelling and Packaging of Substances and Mixtures, Amending and Repealing Directives 67/548/EEC and 1999/45/EC, and Amending Regulation (EC) No 1907/2006, vol. 353:0001, p. 1355 available online at: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2008.

ECHA, 2015. European Chemicals Agency - Guidance on Information Requirements and Chemical Safety Assessment. Chapter R.7a: Endpoint specific guidance. http://echa.europa.eu/web/guest/guidance-documents/guidance-on-reach.

ECHA/RAC, 2013. European Chemicals Agency - Committee for Risk Assessment - Opinion on the ECHA's Draft Review Report on "Evaluation of New Scientific Evidence Concerning DINP and DIDP in Relation to Entry 52 of Annex XVII to Regulation (EC) No 19072006" ECHA-13-r-07-NE ISBN 978-92-9244-001-5, Adopted August 2013.

EFSA, 2005a. Opinion of the scientific panel on food additives, flavourings, processing aids and materials in contact with food (AFC) on a request from the commission related to bis(2-ethylhexyl)phthalate (DEHP) for use in food contact materials. Question N° EFSA-Q-2003-191. Adopted on 23 June 2005 by written procedure. EFSA J. 243, 1–20.

EFSA, 2005b. Opinion of the scientific panel on food additives, flavourings, processing aids and materials in contact with food (AFC) on a request from the commission related to butylbenzylphthalate (BBP) for use in food contact materials. Question N° EFSA-Q-2003-190. Adopted on 23 June 2005 by written procedure. EFSA J. 241, 1–14.

EFSA, 2005c. Opinion of the scientific panel on food additives, flavourings, processing aids and materials in contact with food (AFC) on a request from the commission related to di-butylphthalate (DBP) for use in food contact materials. Question N° EFSA-Q-2003-192. Adopted on 23 June 2005 by written procedure. EFSA J. 242, 1–17.

EFSA, 2005d. Opinion of the scientific panel on food additives, flavourings, processing aids and materials in contact with food (AFC) on a request from the commission related to di-isodecylphthalate (DIDP) for use in food contact materials. Question N° EFSA-Q-2003-195. Adopted on 30 july 2005. EFSA J. 245, 1–14.

EFSA, 2005e. Opinion of the scientific panel on food additives, flavourings, processing aids and materials in contact with food (AFC) on a request from the commission related to di-isononylphthalate (DINP) for use in food contact materials. Question N° EFSA-Q-2003-194. Adopted on 30 july 2005. EFSA J. 244, 1–18.

EU-RAR, 2003. European Union Risk Asessment Report on 1,2-benzenedicarboxylic acid, di-C8-10-branched alkyl esters, C9-rich and di-"isononyl" phthalate (DINP) (CAS Nos: 68515-48-0, 28553-12-0; EINECS Nos: 271-090-9, 249-0979-5). Office for Official Publications of the European Communities, Luxembourg.

EU-RAR, 2008. European Union Risk Asessment Report on bis (2-ethylhexyl) phthalate (DEHP) (CAS No: 117-81-7; EINECS No: 204-211-0). Office for Official Publications of the European Communities, Luxembourg.

Exxon, 1994. Development toxicity study in rats with di-isononyl phthalate (DINP, MRD-92–455). Final Report. Project number 145534. Performed for. Exxon Chemical Company, Houston Texas and Exxon Chemical International, Kraainem, Belgium; Performing Laboratory Exxon Biomedical Sciences Inc., East Millstone, NJ. Date November 30, 1994.

Exxon, 1996a. Reproduction toxicity study in rats with di-isononyl phthalate (DINP, MRD-92–455). Final Report. project number 145535. Performed for. Exxon Chemical Company, Houston Texas and Exxon Chemical Europe, Audergheim, Belgium; Performing Laboratory Exxon Biomedical Sciences Inc., East Millstone, NJ. Date March 8, 1996.

Exxon, 1996b. Two generation reproduction toxicity study in rats with di-isononyl phthalate (DINP, MRD-92–455). Final Report. project number 145535A. Performed for. Exxon Chemical Company, Houston Texas and Exxon Chemical Europe, Audergheim, Belgium; Performing Laboratory Exxon Biomedical Sciences Inc., East Millstone, NJ. Date February 29, 1996.

Field, E.A., Price, C.J., Sleet, R.B., George, J.D., Marr, M.C., Myers, C.B., Schwetz, B.A., Morrissey, R.E., 1993. Developmental toxicity evaluation of diethyl and dimethyl phthalate in rats. Teratology 48, 33–44.

Foster, P.M., Thomas, L.V., Cook, M.W., Gangolli, S.D., 1980. Study of the testicular effects and changes in zinc excretion produced by some n-alkyl phthalates in the rat. Toxicol. Appl. Pharmacol. 54, 392–398.

Furr, J.R., Lambright, C.S., Wilson, V.S., Foster, P.M., Gray Jr., L.E., 2014. A short-term in vivo screen using fetal testosterone production, a key event in the phthalate adverse outcome pathway, to predict disruption of sexual differentiation. Toxicol. Sci. 140, 403–424.

Gray Jr., L.E., Ostby, J., Furr, J., Price, M., Veeramachaneni, D.N., Parks, L., 2000. Perinatal exposure to the phthalates DEHP, BBP, and DINP, but not DEP, DMP, or DOTP, alters sexual differentiation of the male rat. Toxicol. Sci. 58, 350–365.

Gupta, R.C. (Ed.), 2011. Reproductive and Developmental Toxicology. Academic Press as an Imprint of Elsevier, San Diego.

Hellwig, J., Freudenberger, H., Jackh, R., 1997. Differential prenatal toxicity of branched phthalate esters in rats. Food Chem. Toxicol. 35, 501–512.

Hoshino, N., Iwai, M., Okazaki, Y., 2005. A two-generation reproductive toxicity study of dicyclohexyl phthalate in rats. J. Toxicol. Sci. 30, 79–96.

Hushka, L.J., Waterman, S.J., Keller, L.H., Trimmer, G.W., Freeman, J.J., Ambroso, J.L., Nicolich, M., McKee, R.H., 2001. Two-generation reproduction studies in Rats fed di-isodecyl phthalate. Reprod. Toxicol. 15, 153–169.

Jarfelt, K., Dalgaard, M., Hass, U., Borch, J., Jacobsen, H., Ladefoged, O., 2005. Antiandrogenic effects in male rats perinatally exposed to a mixture of di(2-ethylhexyl) phthalate and di(2-ethylhexyl) adipate. Reprod. Toxicol. 19, 505–515.

Juul, A., Almstrup, K., Andersson, A.M., Jensen, T.K., Jorgensen, N., Main, K.M., Rajpert-De Meyts, E., Toppari, J., Skakkebaek, N.E., 2014. Possible fetal determinants of male infertility. Nat. Rev. Endocrinol. 10, 553–562.

Klaassen, C.D. (Ed.), 2013. Casarett and Doull's Toxicology. The Basic Science of Poisons. McGraw-Hill Medical Publishing Division, New York.

Kwack, S.J., Kim, K.B., Kim, H.S., Lee, B.M., 2009. Comparative toxicological evaluation of phthalate diesters and metabolites in Sprague-Dawley male rats for risk assessment. J. Toxicol. Environ. Health A 72, 1446–1454.

Lamb, J.C.t., Chapin, R.E., Teague, J., Lawton, A.D., Reel, J.R., 1987. Reproductive effects of four phthalic acid esters in the mouse. Toxicol. Appl. Pharmacol. 88, 255–269.

Li, L., Bu, T., Su, H., Chen, Z., Liang, Y., Zhang, G., Zhu, D., Shan, Y., Xu, R., Hu, Y., Li, J., Hu, G., Lian, Q., Ge, R.S., 2015. Inutero exposure to diisononyl phthalate caused testicular dysgenesis of rat fetal testis. Toxicol. Lett. 232, 466–474.

Lutter, R., Abbott, L., Becker, R., Borgert, C., Bradley, A., Charnley, G., Dudley, S., Felsot, A., Golden, N., Gray, G., Juberg, D., Mitchell, M., Rachman, N., Rhomberg, L., Solomon, K., Sundlof, S., Willett, K., 2015. Improving weight of evidence approaches to chemical evaluations. Risk Anal. 35, 186–192.

Main, K.M., Skakkebaek, N.E., Virtanen, H.E., Toppari, J., 2010. Genital anomalies in boys and the environment. Best. Pract. Res. Clin. Endocrinol. Metab. 24, 279–289.

Mangelsdorf, I., Buschmann, J., Orthen, B., 2003. Some aspects relating to the evaluation of the effects of chemicals on male fertility. Regul. Toxicol. Pharmacol. 37, 356–369.

Masutomi, N., Shibutani, M., Takagi, H., Uneyama, C., Takahashi, N., Hirose, M., 2003. Impact of dietary exposure to methoxychlor, genistein, or diisononyl phthalate during the perinatal period on the development of the rat endocrine/reproductive systems in later life. Toxicology 192, 149–170.

Myers, J.P., vom Saal, F.S., Akingbemi, B.T., Arizono, K., Belcher, S., Colborn, T., Chahoud, I., Crain, D.A., Farabollini, F., Guillette Jr., L.J., Hassold, T., Ho, S.M., Hunt, P.A., Iguchi, T., Jobling, S., Kanno, J., Laufer, H., Marcus, M., McLachlan, J.A., Nadal, A., Oehlmann, J., Olea, N., Palanza, P., Parmigiani, S., Rubin, B.S., Schoenfelder, G., Sonnenschein, C., Soto, A.M., Talsness, C.E., Taylor, J.A., Vandenberg, L.N., Vandenbergh, J.G., Vogel, S., Watson, C.S., Welshons, W.V., Zoeller, R.T., 2009. Why public health agencies cannot depend on good laboratory practices as a criterion for selecting data: the case of bisphenol A. Environ. Health Perspect. 117, 309–315.

OECD, 2008. Organization for Economic Co-operation and Development – Environment Directorate. Guidance Document on Mammalian Reproductive Toxicity Testing and Assessment. Available at: http://www.oecd.org/document/44/0,3746,en_2649_34377_47787756_1_1_1_1,00.html (Last accessed 18 February 2015).

Ostby, J.B., Horchkiss, A.K., Furr, J.R., Gray, L.E.J., 2001. Investigation of the ability of diisononyl phthalate (DINP) to alter androgen-dependent tissue development in Sprague-Dawley rats. Toxicology 60, 228.

Rhomberg, L., 2015. Hypothesis-based weight of evidence: an approach to assessing causation and its application to regulatory toxicology. Risk Anal. 35, 1114–1124.

Saillenfait, A.M., Gallissot, F., Sabate, J.P., 2009a. Differential developmental toxicities of di-n-hexyl phthalate and dicyclohexyl phthalate administered orally to rats. J. Appl. Toxicol. 29, 510–521.

Saillenfait, A.M., Sabate, J.P., Gallissot, F., 2009b. Effects of in utero exposure to di-n-hexyl phthalate on the reproductive development of the male rat. Reprod. Toxicol. 28, 468–476.

Saillenfait, A.M., Sabate, J.P., Robert, A., Rouiller-Fabre, V., Roudot, A.C., Moison, D., Denis, F., 2013. Dose-dependent alterations in gene expression and testosterone production in fetal rat testis after exposure to di-n-hexyl phthalate. J. Appl.

Toxicol. 33, 1027–1035.
SCCP, 2007. Scientific Committee on Consumer Products: Opinion on Phthtalates in Cosmetic Products. SCCP/1016/06 Adopted on March 21, 2007. Available online at: http://ec.europa.eu/health/ph_risk/committees/04_sccp/docs/sccp_o_106.pdf.
Schreider, J., Barrow, C., Birchfield, N., Dearfield, K., Devlin, D., Henry, S., Kramer, M., Schappelle, S., Solomon, K., Weed, D.L., Embry, M.R., 2010. Enhancing the credibility of decisions based on scientific conclusions: transparency is imperative. Toxicol. Sci. 116, 5–7.
Scott, H.M., Hutchison, G.R., Mahood, I.K., Hallmark, N., Welsh, M., De Gendt, K., Verhoeven, G., O'Shaughnessy, P., Sharpe, R.M., 2007. Role of androgens in fetal testis development and dysgenesis. Endocrinology 148, 2027–2036.
Sharpe, R.M., Skakkebaek, N.E., 2008. Testicular dysgenesis syndrome: mechanistic insights and potential new downstream effects. Fertil. Steril. 89 e33–8.
Spade, D.J., Hall, S.J., Wilson, S., Boekelheide, K., 2015. Di-n-butyl phthalate induces multinucleated germ cells in the rat fetal testis through a nonproliferative mechanism. Biol. Reprod. http://dx.doi.org/10.1095/biolreprod.115.131615 published ahead of print September 23, 2015.
Taxvig, C., Hass, U., Axelstad, M., Dalgaard, M., Boberg, J., Andeasen, H.R., Vinggaard, A.M., 2007. Endocrine-disrupting activities in vivo of the fungicides tebuconazole and epoxiconazole. Toxicol. Sci. 100, 464–473.
US-EPA, 2005. Environmental Protection Agency. Guidelines for Carcinogen Risk Assessment. EPA/630/P-03/001B. EPA, Washington, DC. Available at: http://www.epa.gov/raf/publications/pdfs/CANCER_GUIDELINES_FINAL_3-25-05.pdf.
Van Der Kraak, G.J., Hosmer, A.J., Hanson, M.L., Kloas, W., Solomon, K.R., 2014. Effects of atrazine in fish, Amphibians, and reptiles: an analysis based on quantitative weight of evidence. Crit. Rev. Toxicol. 44, 1–66.
Waterman, S.J., Ambroso, J.L., Keller, L.H., Trimmer, G.W., Nikiforov, A.I., Harris, S.B., 1999. Developmental toxicity of di-isodecyl and di-isononyl phthalates in rats. Reprod. Toxicol. 13, 131–136.
Waterman, S.J., Keller, L.H., Trimmer, G.W., Freeman, J.J., Nikiforov, A.I., Harris, S.B., Nicolich, M.J., McKee, R.H., 2000. Two-generation reproduction study in rats given di-isononyl phthalate in the diet. Reprod. Toxicol. 14, 21–36.
Weed, D.L., 2005. Weight of evidence: a review of concept and methods. Risk Anal. 25, 1545–1557.
Yamasaki, K., Okuda, H., Takeuchi, T., Minobe, Y., 2009. Effects of in utero through lactational exposure to dicyclohexyl phthalate and p,p'-DDE in Sprague-Dawley rats. Toxicol. Lett. 189, 14–20.