Chemosphere 188 (2017) 174–181

Contents lists available at ScienceDirect

Chemosphere

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

Review

Systematic review and meta-analysis of early life exposure to di(2-ethylhexyl) phthalate and obesity related outcomes in rodents

Pim Nicolaas Hubertus Wassenaar*, Juliette Legler [1]

Institute for Environmental Studies, VU University Amsterdam, Amsterdam, The Netherlands

HIGHLIGHTS

- Early life exposure to DEHP was significantly associated with increased fat weight.
- A non-significant negative association was estimated for body weight.
- There was substantial heterogeneity across studies.
- Reported information was insufficient to assess the risk of bias for most studies.
- More data is necessary to strengthen the evidence base of the obesogenic effects.

ARTICLE INFO

Article history:
Received 15 July 2017
Received in revised form
28 August 2017
Accepted 30 August 2017
Available online 31 August 2017

Handling editor: A. Gies

Keywords:
Systematic review
Obesity
Early life exposure
Endocrine disrupting chemicals
di(2-ethylhexyl) phthalate
mono(2-ethylhexyl) phthalate

ABSTRACT

Background: It has been suggested that the plasticizer di(2-ethylhexyl) phthalate (DEHP) exerts obesogenic effects after pre- or perinatal exposure.
Objective: A systematic review with meta-analyses was conducted of early life exposure to DEHP, or its biologically active metabolite mono(2-ethylhexyl) phthalate (MEHP), on the obesity related outcome measures body weight, fat (pad) weight, triglycerides, free fatty acids and leptin in experimental rodent studies.
Methods: The applied methodology was pre-specified in a rigorous protocol. Relevant articles were identified using PubMed and EMBASE and meta-analyses were performed using mean differences (MD) and random effects model when at least five studies could be included per outcome measure. Risk of bias and the quality of evidence was assessed using established methodologies.
Results: Overall, 31 studies could be included and meta-analyses could be performed for body weight and fat weight. Early life exposure to DEHP was significantly associated with increased fat weight (MD = 0.02; 95% CI: 0.00, 0.03), while a non-significant association was estimated for body weight (MD = −0.14; 95% CI: −0.32, 0.04). There was substantial heterogeneity across studies and the information was insufficient to assess the risk of bias for most studies. No meta-analyses could be conducted for other outcome measures, because too few studies were available.
Conclusions: The results of this systematic review indicate that early life exposure to DEHP is potentially associated with increased adiposity in rodents. More data is needed to strengthen the evidence base.

Contents

* Corresponding author. Institute for Environment Studies, VU University Amsterdam, De Boelelaan 1087, 1081 HV, Amsterdam, The Netherlands.
E-mail address: pim.wassenaar@gmail.com (P.N.H. Wassenaar).
[1] Current address: Institute of Environment, Health and Societies, College of Health and Life Sciences, Brunel University London, Uxbridge, United Kingdom.

http://dx.doi.org/10.1016/j.chemosphere.2017.08.165

1. Introduction

Di(2-ethylhexyl) phthalate (DEHP) is used in many plastic products as plasticizer, including in packaging material, flooring, cables and medical products (ECHA, 2009). Its use has been subject of discussion as DEHP is ubiquitous in the environment (Erythropel et al., 2014) and has been identified as a substances with reproductive toxicity properties (ECHA, 2008; NTP, 2006). In addition, within Europe, DEHP has also been identified as endocrine disrupting chemical (EDC) in humans and the environment, and its use has been restricted in many products (ECHA, 2017, 2014).

In addition to the effects of DEHP on reproductive endpoints, DEHP has putative obesogenic properties. Obesogens are chemicals that can inappropriately regulate lipid metabolism via hormonal pathways, and thereby stimulate lipid accumulation and adipogenesis or could result in higher susceptibility to develop obesity (Grün and Blumberg, 2006). Early life exposure has been suggested to be the critical exposure period for the occurrence of these obesogenic effects, and is of concern as DEHP is found in urine samples of children in the US and EU (Casas et al., 2013; DEMOCOPHES, 2013; Zota et al., 2014), and is identified in maternal blood, cord blood and meconium of mother-newborn pairs (Li et al., 2013). Obesogenic effects of DEHP have been demonstrated in multiple studies including in *in vitro* and animal experiments as well as in epidemiological studies (Gore et al., 2015). It has been suggested that DEHP exerts its endocrine disrupting effects mainly via its primary metabolite mono(2-ethylhexyl) phthalate (MEHP), through activation of the nuclear peroxisome proliferator-activated receptors (PPAR) as well as PPAR mediated anti-androgenic effects (Desvergne et al., 2009; Kim and Park, 2014). Of these receptors, PPARα and −γ play a pivotal role in fatty acid and lipid metabolism and adipogenesis (Desvergne et al., 2009). In addition, potential thyroid and estrogenic disrupting have been suggested as a mode of action of DEHP (ECHA, 2014; Kim and Park, 2014).

In order to provide a more rigorous evaluation of the existing evidence we conducted a systematic review on the effects of DEHP exposure on obesity development. In this review, we investigated the effects of pre- or perinatal exposure of rodents to DEHP and MEHP on the following obesity related outcome measures: body weight, fat (pad) weight, triglycerides, free fatty acids (FFA), and leptin (Table 1). We determined the quality of the studies, rated the confidence of the evidence using established methodologies and, where possible, conducted meta-analyses.

2. Methods

The methodology of this systematic review was pre-specified in a protocol, following guidelines of the Systematic Review Centre for Laboratory animal Experimentation (SYRCLE; de Vries et al., 2015). This protocol has been published on the SYRCLE Website in 2015 and is provided in the Appendix. The applied methodology is similar to the methodology as described by Wassenaar et al. (In Press).

2.1. Search strategy and selection of papers

A comprehensive search strategy was applied on September 21, 2015 in order to identify relevant articles in the databases MEDLINE via PubMed and EMBASE (Table A.1). In addition, the reference lists of the included articles and of relevant reviews were screened manually in order to identify potentially relevant new articles.

Study selection consisted of two screening phases, including a title and abstract screening and a full text screening. Studies were selected for full text screening when they met all the inclusion criteria: 1) original full paper which presented unique data; 2) exposure to DEHP or MEHP; 3) obesity related article or at least one of the outcome measures was examined (body weight, fat pad weights, triglyceride levels, FFA levels or leptin levels); 4)

Table 1
PECO statement (Population, Exposure, Comparator and Outcomes).

Population	Experimental rodent studies
Exposure	Early life exposure to di(2-ethylhexyl) phthalate or mono(2-ethylhexyl) phthalate (during gestation and/or lactation up to postnatal day 21)
Comparator	Animals exposed to vehicle-only treatment
Outcomes	Body weight, fat (pad) weights, triglyceride levels, free fatty acids levels and leptin levels

experimental rodent study; and 5) perinatal exposure via maternal or direct pup exposure (during gestation and/or lactation up to postnatal day (PND) 21). In case of doubt, articles were also analyzed based on their full text. During full text screening studies were excluded if they met one of the following criteria: 1) not an original paper; 2) exposure to a different chemical than DEHP or MEHP; 3) not disease or outcome of interest (no obesity related outcome); 4) not a rodent study; 5) not perinatal exposure (paternal exposure, exposure after PND21 and measurements in unborn fetuses were excluded); 6) outcomes not measured in F1 generation; 7) unhealthy or genetically altered rodents (data measured after ovariectomy were seen as unhealthy and were not extracted); 8) outcomes were measured after diet was altered to high fat diet during follow-up. In addition, selection was restricted to English language articles and to articles which were freely accessible via available sources. The screening phases were conducted by one reviewer (PW) and in case of doubt a second reviewer (JL) was consulted.

2.2. Study characteristics and data extraction

From the included studies we extracted bibliographic data, animal model characteristics, exposure characteristics, study design characteristics and relevant outcome measures (i.e. body weight, fat (pad) weights, triglyceride levels, FFA levels and/or leptin levels). From each study, we considered each analysis of a specific outcome measure with a specific dose and/or gender as a separate individual comparison. In addition, analyses in different fat pads (even when derived from the same animal) and analyses with different time windows of exposure were considered as separate comparisons. Consequently, multiple comparisons could have been included from one study. Outcome data were extracted and collected for each individual comparison as mean, SD and number of animals per group as described by Wassenaar et al. (In Press). When outcomes were measured at different time points, the time point with greatest efficacy was used (i.e. the time point with the strongest association with the outcome). The time point of greatest efficacy was selected over the other time point(s) when the absolute difference between the mean of the exposure and control group divided by the sum of the SDs was highest. By using the absolute difference the direction of the effect was not considered in the selection of the time point of greatest efficacy. When data were not sufficiently reported we contacted authors by e-mail for clarification. When no response was received after repeated contact within three weeks, the data were omitted from the analyses. Data extraction was conducted by one reviewer (PW) and in case of doubt a second reviewer was consulted (JL).

2.3. Risk of bias and methodological quality assessment

The risk of bias in the included studies was assessed similarly as described by Wassenaar et al. (In Press), using the SYRCLE's Risk of Bias (RoB) tool which is specifically designed for animal studies (Hooijmans et al., 2014). Briefly, the ten items in the RoB tool were scored with “yes” indicating low risk of bias, “no” indicating high risk of bias or “?/unclear” indicating that the item was not reported and therefore the risk of bias was unknown. In addition, four items were added to check the methodological quality. Two of these items examined potential litter effects, focusing on the applied statistical units (i.e. litter or offspring) and effects on litter size. These items were scored with “yes”, “no” or “?/unclear”. Furthermore, two overall study quality indicators focusing on the reporting of any randomization or any blinding were included. These items were scored with “yes” when reported or “no” when not reported. The RoB and methodological quality assessment were conducted by one reviewer (PW) and in case of doubt a second reviewer (JL) was consulted.

2.4. Data synthesis and statistical analysis

Meta-analyses were performed using Review Manager v5.3 (The Cochrane Collaboration, 2014) when at least five studies reported on a specific outcome measure, which was the case for body weight and fat weight. When less than five studies could be included, data were presented narratively, which was the case for triglyceride and leptin. Meta-analyses were conducted using mean differences (MD; the mean of the experimental group minus the mean of the control group) and random effect models in order to account for the anticipated heterogeneity. Heterogeneity was assessed using I-square (I^2) and the significance level of the meta-analyses was set at $p < 0.05$.

Furthermore, subgroup meta-analyses were performed in order to assess the influence of variables where at least three studies could be included per subgroup. The assessed subgroup variables include: animal species (rats or mice), strains, gender (male or female), time window of exposure (perinatal exposure: during gestation and lactation period, prenatal exposure: during gestation only, or postnatal exposure: during lactation period only), dosage of treatment (below or above the tolerable daily intake (TDI) of 50 μg/kg/d; EFSA, 2005), route of exposure (gavage, oral, diet, drinking water or subcutaneous injections), time of outcome measurement (before or after PND21, the normal weaning period) and frequency of exposure (daily or constant exposure, for example via constant availability of DEHP or MEHP in drinking water or diet). Some studies provide birth weight data for mixed genders, whereas at later time points, sex-specific body weight data is provided. When available, we included only the sex-specific data as these are considered to be more informative (i.e. they provide data on sex-specific effects as well as information on later time points). For the outcome measure fat weight, subgroup analyses were conducted to analyze the effects on different fat pads. All subgroup differences were analyzed with a test for subgroup differences in Review Manager. In order to correct for multiple testing, the significance levels of the subgroup analyses were adjusted to $p < 0.01$.

2.5. Sensitivity analyses

Three sensitivity analyses were performed in order to evaluate the robustness of the results. First, the impact of the latest time point on interpreting study results was assessed by selecting the latest measured time point instead of the time point with greatest efficacy. Second, a sensitivity analysis was performed by excluding studies of potential high bias and poor quality, i.e. excluding studies that scored not a single “yes” score on items 1–12 of the risk of bias and the methodological quality indicators. Third, we conducted a sensitivity analysis by excluding studies which were not free of potential litter effects based on the two included methodological quality items (items 13–14).

2.6. Confidence rating

We rated the quality of evidence of the outcomes of this systematic review according to the confidence rating methodology of the Office of Health Assessment and Translation (OHAT; NTP, 2015) as described in Wassenaar et al. (In Press). Briefly, an initial high confidence can be downgraded to a moderate, low or very low confidence based on assessment of five factors (i.e. risk of bias, unexplained inconsistency, indirectness, imprecision and publication bias). In addition, four factors were assessed that can increase the confidence rating (i.e. large magnitude of effect, dose response,

plausible confounding and consistency across study designs). Similarly to Wassenaar et al. (In Press), the factor "large magnitude of effect" was not assessed as it is difficult to interpret a large effect because a relatively small effect could have major public health consequences. Furthermore, the factor "plausible confounding" was not assessed as this factor primarily applies to observational studies (NTP, 2015). Publication bias was assessed with funnel plots when at least 10 studies could be included, otherwise only lag time for negative studies and conflict of interest sections were investigated.

3. Results

3.1. Study selection and characteristics

A flow chart of the study selection process is provided in Fig. 1. Using the comprehensive search strategies, 2535 unique articles were identified from PubMed and EMBASE and after screening a total of 31 articles could be included in the systematic review. Of these articles, 30 reported effects on body weight, 6 on fat weight, 3 on triglyceride levels, 1 on leptin levels and none on FFA levels.

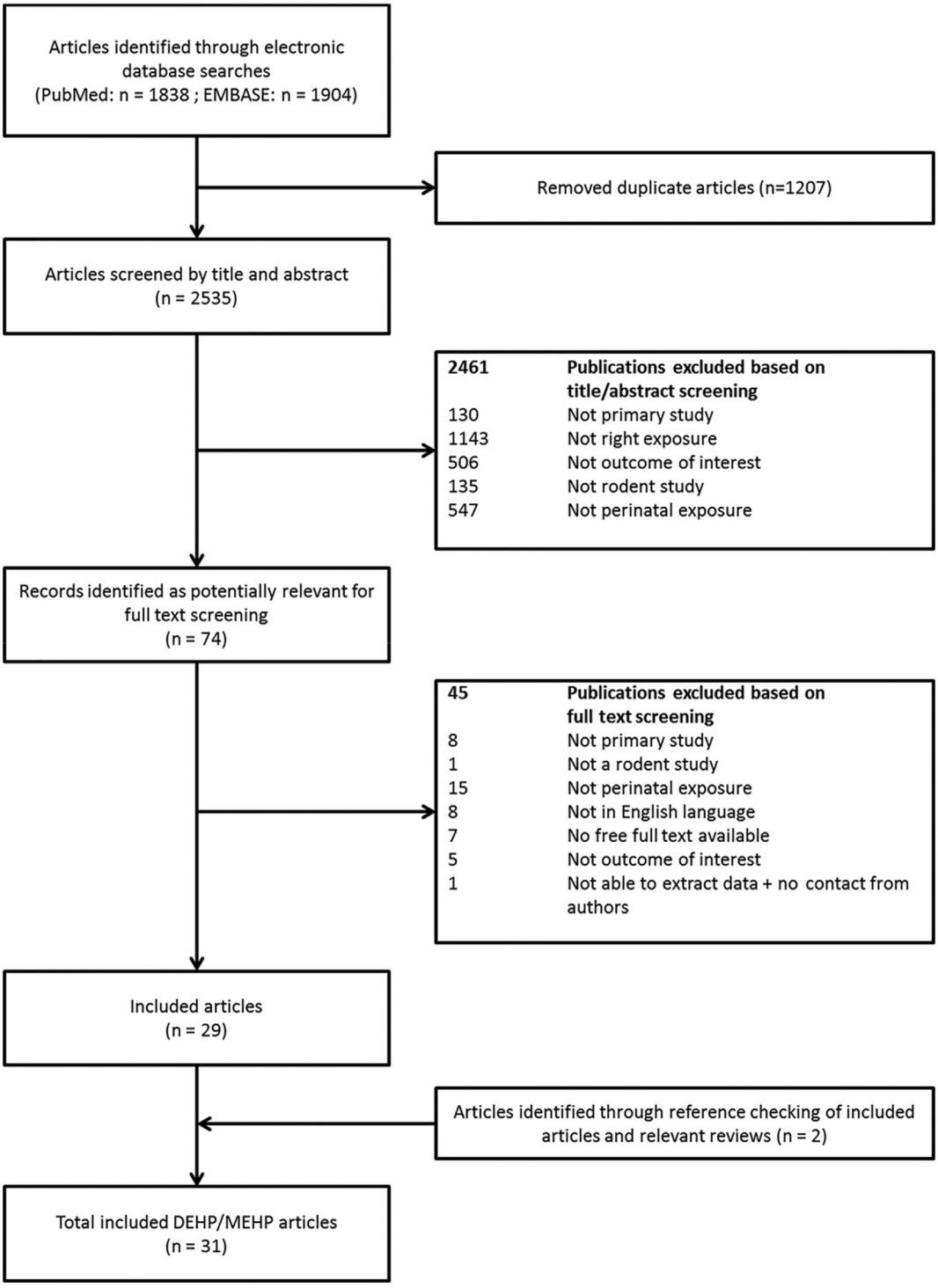

Fig. 1. Flow chart of study selection process.

These outcome measures include 149, 30, 15, and 4 independent comparisons, respectively, totaling 198 independent comparisons. Detailed characteristics of all included studies are provided in Table A.2.

3.2. Risk of bias and methodological quality assessment

The main observation from the risk of bias and methodological quality assessment is the many "unclear" scores, indicating that most items were not sufficiently reported, resulting in an unknown risk of bias (Fig. 2). As a result, most items of the risk of bias assessment could not be judged. The individual scores of the RoB tool and the methodological quality indicators of each included study are provided in Table A.3. Items which were more frequently recorded include baseline similarities (52% had a low risk of bias; Q2), incomplete outcome data (three studies had a low risk of bias and one study a high risk of bias; Q8), and selective outcome reporting (two studies had a high risk of bias; Q9). Furthermore, the methodological quality indicators indicate that 71% of the studies report any randomization, whereas any blinding was only reported in 32% of the studies. Assessment of litter effects revealed that litter was used as statistical unit in 32% of the studies, the offspring in 29% of the studies and in 32% of the studies it was unclear whether offspring or litter were used. In addition, two studies received a "yes/no" score, as the litter was used as statistical unit in measurements before PND21, whereas after PND21 the offspring were used. Litter size was affected after exposure in 13% of the studies.

3.3. Effects of DEHP and MEHP on obesity related outcomes

3.3.1. Body weight

Out of 30 studies, a total of 149 comparisons investigating the effects of DEHP and MEHP on body weight (in grams) could be included in the meta-analyses. Overall, DEHP and MEHP exposure were associated with a non-significant decrease in body weight (MD = −0.14; 95% CI: −0.32, 0.04; Table 2), with substantial heterogeneity among studies (I^2 = 92%). In Figs. A.1–A.9, forest plots are provided of the overall and subgroup analyses showing the effect estimates of DEHP and MEHP exposure on body weight. Subgroup analyses showed that heterogeneity was very high for all estimates (i.e. I^2 > 75%). A significant negative association was estimated for rats (MD = −0.36; 95% CI: −0.63, −0.10), while a non-significant positive association was estimated for mice (MD = 0.14; 95% CI: −0.14, 0.42; p-value for subgroup differences = 0.01; Table 2). Significant negative associations were estimated for CD-1 mice and Sprague-Dawley rats and null or non-significant positive associations for other strains (p-value for subgroup differences = 0.0005). Associations did not vary between gender (p-value for subgroup differences = 0.12) and frequency of exposure (p-value for subgroup differences = 0.05). Postnatal exposure was associated with a significant decreased body weight, whereas null and non-significant positive associations were estimated for perinatal and prenatal exposure, respectively (p-value for subgroup differences < 0.00001). Furthermore, a significant negative association was estimated for oral exposure, mainly through pipette feeding, while for gavage and diet exposure null or non-significant

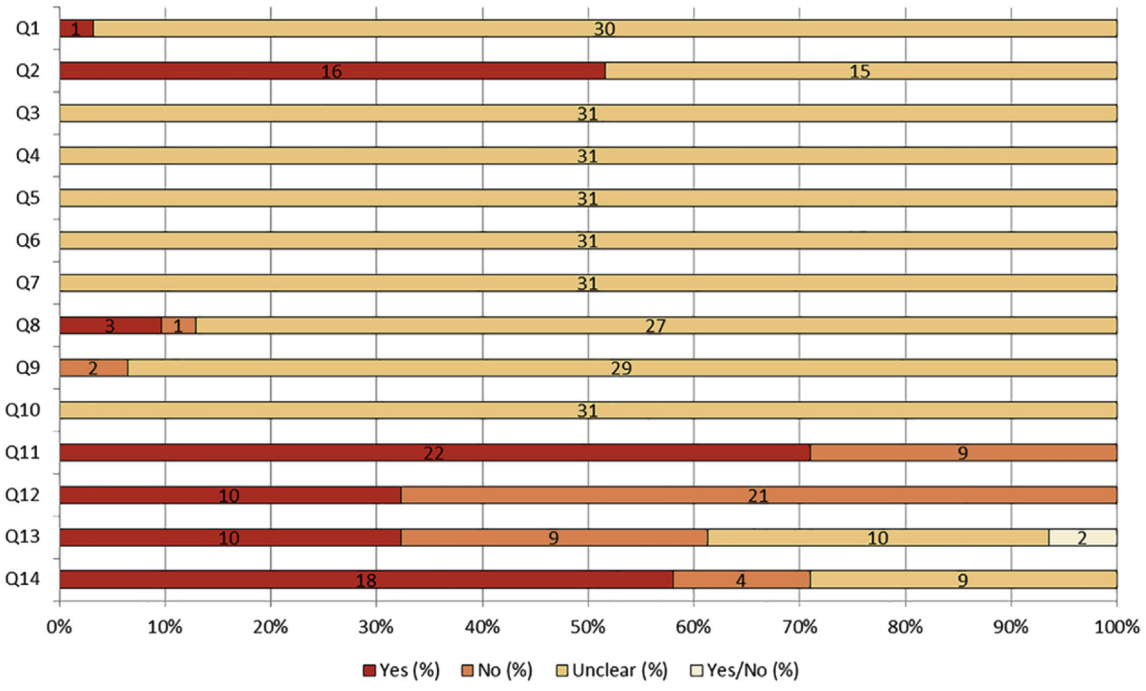

Fig. 2. Results of the risk of bias and methodological quality indicators for all included studies. Notes: The items in the SYRCLE Risk of Bias assessment (Q1-Q10) were scored with "yes" indicating low risk of bias, "no" indicating high risk of bias or "unclear" indicating that the item was not reported, resulting in an unknown risk of bias (Hooijmans et al., 2014). Q1-Q3 considers selection bias; Q4-Q5 performance bias; Q6-Q7 detection bias; Q8 attrition bias; Q9 reporting bias; and Q10 other biases. The overall study quality indicators (Q11-Q12) were scored with "yes" when reported or "no" when not reported. The methodological quality indicators focusing on potential intralitter correlation (Q13-Q14) were scored with "yes", "no" or "unclear". Q = question; Q1 = Was the allocation sequence adequately generated and applied?; Q2 = Were the groups similar at baseline or were they adjusted for confounders in the analysis?; Q3 = Was the allocation to the different groups adequately concealed?; Q4 = Were the animals randomly housed during the experiment?; Q5 = Were the caregivers and/or investigators blinded from knowledge which intervention each animal received during the experiment?; Q6 = Were animals selected at random for outcome assessment?; Q7 = Was the outcome assessor blinded?; Q8 = Were incomplete outcome data adequately addressed?; Q9 = Are reports of the study free of selective outcome reporting?; Q10 = Was the study apparently free of other problems that could result in high risk of bias?; Q11 = Was it stated that the experiment was randomized at any level?; Q12 = Was it stated that the experiment was blinded at any level?; Q13 = Was intralitter correlation controlled for by using the litter as statistical unit (instead of offspring)?; Q14 = Was the study free of potential intralitter correlation caused by effects on litter size?.

Table 2
Effects of early life exposure to DEHP and MEHP on body weight and fat weight from random effect meta-analyses.

Analysis	Subgroups	MD (gram) ± 95% CI	I^2	# studies (# comparisons)	Test for subgroup differences
Body weight					
Overall		−0.14 [−0.32, 0.04]	92%	30 (149)	
Species	Rats	−0.36 [−0.63, −0.10]	81%	22 (119)	p = 0.01
	Mice	0.14 [−0.14, 0.42]	97%	8 (30)	
Strain	CD-1 (mice)	−0.40 [−0.61, −0.19]	93%	4 (13)	p = 0.0005
	Sprague-Dawley (r)	−1.11 [−1.69, −0.54]	84%	12 (45)	
	Wistar (rats)	−0.02 [−0.31, 0.26]	78%	9 (72)	
	Others [a]	0.68 [−0.13, 1.50]	98%	5 (19)	
Gender [b]	Males	−0.27 [−0.58, 0.03]	86%	25 (85)	p = 0.12
	Females	0.08 [−0.25, 0.41]	94%	15 (58)	
Time window of exposure	Perinatal	0.03 [−0.28, 0.35]	93%	17 (99)	p < 0.00001
	Prenatal	0.12 [−0.02, 0.27]	68%	8 (27)	
	Postnatal	−1.71 [−2.43, −1.00]	88%	5 (23)	
Frequency of exposure [c]	Daily	−0.28 [−0.52, −0.05]	82%	24 (126)	p = 0.05
	Constant	0.20 [−0.23, 0.63]	99%	4 (15)	
Route of exposure [d]	Gavage	0.00 [−0.22, 0.23]	77%	20 (109)	p = 0.002
	Oral	−1.87 [−2.96, −0.79]	89%	5 (23)	
	Diet	0.20 [−0.23, 0.63]	99%	4 (15)	
Dose of exposure	<50 μg/kg	0.23 [−0.37, 0.83]	96%	9 (17)	p = 0.17
	>50 μg/kg	−0.21 [−0.40, −0.02]	91%	29 (132)	
Time of outcome measure	PND0 - PND21	−0.07 [−0.25, 0.11]	95%	17 (77)	p = 0.13
	> PND21	−0.80 [−1.74, 0.13]	76%	19 (72)	
Fat weight					
Overall		0.02 [0.00, 0.03]	92%	6 (30)	
Gender	Males	0.03 [−0.00, 0.06]	94%	6 (16)	p = 0.26
	Females	0.01 [−0.01, 0.02]	87%	5 (14)	
Dose of exposure	<50 μg/kg	0.05 [−0.00, 0.10]	92%	3 (8)	p = 0.10
	>50 μg/kg	0.00 [−0.01, 0.02]	90%	6 (22)	

Notes: Effect sizes are expressed in mean differences (MD) with 95% confidence interval calculated using random effects model. From each study, we considered each analysis with a specific dose, gender and/or time window of exposure as a separate individual comparison, as well as analyses of different fat pads (even when derived from the same animal). I^2 is a measure of heterogeneity. Positive MDs represent an increase in the outcome measure after exposure. Negative MDs represent a decrease in the outcome measure after exposure. Test for subgroup differences was conducted using Review Manager (v5.3).

[a] Includes C57BL/6J, sv/129 and C3H/N mice and Long-Evans rats.

[b] Only two studies examined mixed genders and therefore these studies were excluded for this subgroup analysis (Table A.2).

[c] Two studies were excluded as exposure was not on a constant or daily basis (Table A.2).

[d] One study was excluded examining subcutaneous exposure (Table A.2).

positive associations were estimated (p-value for subgroup differences = 0.002). In addition, a non-significant positive association was estimated for an exposure dose below 50 μg/kg/d, and a non-significant negative association for an exposure dose above 50 μg/kg/d (p-value for subgroup differences = 0.17). Also, no difference in association was estimated for time of outcome measurements (p-value for subgroup differences = 0.13).

3.3.2. Fat weight

In total, 6 studies consisting of 30 comparisons reported effects on fat weight (in grams) and could be included. Overall, DEHP and MEHP exposure was associated with an increased fat weight (MD = 0.02; 95% CI = 0.00, 0.03; Table 2) with substantial heterogeneity among the studies (I^2 = 92%). In Figs. A.10–A.12, forest plots are provided of the overall and subgroup analyses showing the effect estimates of DEHP and MEHP exposure on fat weight. Also for fat weight, subgroup analysis showed that heterogeneity was very high for all estimates (i.e. I^2 > 75%). No differences in association was estimated for gender (p-value for subgroup differences = 0.26; Table 2). Also for dose of exposure no difference in association were estimated (p-value for subgroup differences = 0.10) with a non-significant positive association for doses below 50 μg/kg/d (MD = 0.05; 95% CI: −0.00, 0.10) and a null association for doses above 50 μg/kg/d (MD = 0.00; 95% CI: −0.01, 0.02). No other subgroup analyses could be conducted because less than three studies could be included per subgroup.

3.3.3. Triglycerides

Three studies described the effects of DEHP and MEHP exposure on triglyceride levels. Two of these studies examined effects of perinatal exposure to DEHP or MEHP in mice on serum triglycerides. Perinatal DEHP exposure (0.25 mg/kg/d) from gestational day (GD) 12 to PND7 resulted in increased triglyceride levels in males and females on PND56 (Hao et al., 2013). Perinatal MEHP exposure (from GD12 to PND7) resulted in elevated triglyceride levels in male mice offspring at 0.05 mg/kg/d on PND56, whereas no effect was observed at higher concentrations of 0.25 and 0.5 mg/kg/d. In addition, no effect was observed in female mice (Hao et al., 2012). Another study investigated the effect of postnatal DEHP exposure on plasma triglyceride levels in rats. No effect on triglyceride levels was observed after exposure in the period PND6-PND10, PND14-PND18 or PND16-PND20 at concentrations of 10, 100 or 1000 mg/kg/d. Outcomes were measured 24 h after the last dose was administered (Dostal et al., 1987).

3.3.4. Leptin

Only 1 study, consisting of 4 comparisons, examined the effects of prenatal DEHP exposure on leptin levels (Campioli et al., 2014). Effects of four different concentrations ranging from 1 to 300 mg/kg/d were analyzed in Sprague-Dawley rats on PND60. Within this study no effects on leptin levels were observed for all exposure concentrations.

3.4. Sensitivity analyses

In all three sensitivity analyses the overall effect of body weight changed from a non-significant negative effect to a significant negative effect: MD = −0.44 (95% CI:−0.62, −0.26) after the

sensitivity analysis using data from the latest measured time point, MD = −0.19 (95% CI: −0.37, −0.00) after the sensitivity analysis excluding studies of potential high bias, and MD = −0.46 (95% CI: −0.89, −0.02) after the sensitivity analysis excluding studies which were not free of potential litter effects. The overall effect on fat weight did not change when the latest measured time points were used (MD = 0.03; 95% CI: 0.01, 0.05). The other sensitivity analyses could not be conducted for fat weight, as too many studies had to be excluded. All comparisons for which the individual effect estimate had to be changed to the latest measured time point, and all studies which had to be excluded for the other sensitivity analyses are indicated in Table A.2 and Table A.3, respectively.

3.5. Confidence rating

The quality of evidence of body weight and fat weight were both downgraded because of serious concerns for risk of bias due to the many unclear scores (Table 3). In addition, the confidence of both outcome measures was downgraded with respect to unexplained inconsistency because substantial heterogeneity (i.e. $I^2 > 75\%$), varying point estimates and a minimal CI overlap were observed. Publication bias was judged as undetected, based on the funnel plots (Fig. A.13), conflict of interest sections (Table A.3) and the absence of a lag-time for "negative" studies. Consequently, the quality of evidence of body weight and fat weight was rated as low.

4. Discussion

The results of this systematic review of the rodent literature indicate that early life exposure to DEHP is positively associated with fat weight, while a non-significant negative association was estimated with body weight. Within the data there was substantial heterogeneity across studies and the reported information was insufficient to assess the risk of bias for most studies.

Subgroup analysis of fat weight did not indicate an effect of dose, suggesting that doses below the current TDI level of 50 μg/kg/d in Europe were also associated with increased fat weight. The current TDI has been allocated by the European Food and Safety Authority (EFSA) based on a no observed adverse effect level (NOAEL) of 5 mg/kg/d for testicular and developmental toxicity (EFSA, 2005). Within the US, the US Environmental Protection Agency (EPA) has set the reference dose (RfD; equivalent to the TDI in Europe) at 20 μg/kg/d based on a lowest observed adverse effect level (LOAEL) of 19 mg/kg/d for increased liver weight (US EPA, 1987). Additional data is necessary to assess whether effects on fat weight occur at these dose levels and whether the current limit values are sufficiently safe.

Assessed subgroup variables indicate varying associations of body weight with species and strains, as well as with time window and route of exposure. However, given the limitations of the available data, these subgroup results should be interpreted with caution.

Due to the lack of data on triglycerides, leptin and FFAs no meta-analyses could be conducted for these outcome measures. Available data indicate potential sex-specific effects on triglyceride levels as well as effects at low exposure doses, however, additional data is necessary to substantiate these effects.

4.1. Strengths and limitations

The methodology applied in this systematic review is based on a rigorous protocol following the SYRCLE approach, which has been specifically designed to evaluate animal studies. This systematic review is limited by the poor reporting quality of included animal studies as well as on the data availability on the effects of early life exposure to DEHP and MEHP on obesity related outcome measures. Consequently, only meta-analyses could have been conducted on body weight and fat weight. The lack of sufficient reporting quality is notable and it is recommended to improve the reporting quality by using checklists when submitting manuscripts, like the ARRIVE guidelines (Kilkenny et al., 2010) and the Gold Standard Publication Checklist (Hooijmans et al., 2010).

Within the data there was a substantial amount of heterogeneity, which was not notably reduced after subgroup analyses. This indicates that variations in the design and quality of included studies are the main sources of heterogeneity in this systematic review. In order to account for the heterogeneity, we applied the random effect model.

Within this systematic review, one reviewer conducted the screening phases, data extraction and study quality assessment with rigorous consultation between reviewers. However, in general, the use of two independent reviewers is preferred as the use of one reviewer might result in more errors (Buscemi et al., 2006).

Furthermore, this review was restricted to English language articles and to articles which were freely accessible via available sources. Due to this restriction, 8 non-English articles were excluded and 7 articles were not freely available. The impact of these restrictions may be considered as limited, as these studies mainly described effects on body weight according to their abstracts, for which already 30 studies were included. In addition, this study only focused on rodents and did not consider other species. Rodents are specifically relevant for risk assessment purposes, however, future research on additional species would be a valuable addition. Though the developmental time period is critically sensitive to both nutritional and environmental influences that can affect the etiology of obesity (Heindel and Schug, 2013), we did not

Table 3
Quality of the evidence of the overall effects of DEHP/MEHP on the investigated obesity related outcome measures using the OHAT confidence rating methodology (NTP, 2015).

Outcome measure	Body of Evidence (animal studies)	Factors for downgrading					Factors for upgrading				Final Confidence Rating
		Risk of Bias	Unexplained Inconsistency	Indirectness	Imprecision	Publication Bias	Magnitude [f]	Dose Response	Residual Confounding [f]	Consistency Species	
Body weight	Initial high confidence (30 studies)	↓[a]	↓[c]	–	––[d]	––[e]	–	–	–	–	Low
Fat weight	Initial high confidence (6 studies)	↓[b]	↓[c]	–	–	––[e]	–	–	–	–	Low

Notes: "–" = no concern, or not present; ↓ = serious concern; ↑ = sufficient to upgrade evidence.
[a] Serious concern because of the many "unclear" scores and a change in direction of the association after sensitivity analyses.
[b] Serious concern because of the many "unclear" scores.
[c] Serious concern because of varying point estimates, minimal or no overlap of confidence intervals between studies and substantial heterogeneity ($I^2 > 75\%$).
[d] Body of evidence is already downgraded for unexplained inconsistency and additional downgrading for imprecision is not considered appropriate (NTP, 2015).
[e] No strongly suspected publication bias observed.
[f] The factors "large effect magnitude" and "residual confounding" were not assessed in this study and consequently not used to upgrade the evidence.

examine the effects of diet in our review. We excluded high-fat diet challenges as inclusion of additional variables was associated with an increased complexity.

5. Conclusions

In conclusion, this systematic review indicates that early life exposure to DEHP is associated with increased fat weight in rodents, while a non-significant negative association was estimated with body weight. It should be noted that there was substantial heterogeneity across studies, and that information was insufficient to assess the risk of bias for most studies. More data is necessary to strengthen the evidence base of the obesogenic effects of DEHP.

Conflicts of interest

None.

Acknowledgements

This project received funding from the European Community's Seventh Framework Programme [FP7/2007–2013] under grant agreement OBELIX n° 227391 and the Netherlands Organization for Scientific Research (NWO-VIDI 864.09.005).

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.chemosphere.2017.08.165.

References

Buscemi, N., Hartling, L., Vandermeer, B., Tjosvold, L., Klassen, T.P., 2006. Single data extraction generated more errors than double data extraction in systematic reviews. J. Clin. Epidemiol. 59, 697–703. http://dx.doi.org/10.1016/j.jclinepi.2005.11.010.

Campioli, E., Martinez-Arguelles, D.B., Papadopoulos, V., 2014. In utero exposure to the endocrine disruptor di-(2-ethylhexyl) phthalate promotes local adipose and systemic inflammation in adult male offspring. Nutr. Diabetes 4, e115. http://dx.doi.org/10.1038/nutd.2014.13.

Casas, M., Chevrier, C., Hond, E., Den Fernandez, M.F., Pierik, F., Philippat, C., Slama, R., Toft, G., Vandentorren, S., Wilhelm, M., Vrijheid, M., 2013. Exposure to brominated flame retardants, perfluorinated compounds, phthalates and phenols in European birth cohorts: ENRIECO evaluation, first human biomonitoring results, and recommendations. Int. J. Hyg. Environ. Health. http://dx.doi.org/10.1016/j.ijheh.2012.05.009.

de Vries, R.B.M., Hooijmans, C.R., Langendam, M.W., Leenaars, M., Ritskes-Hoitinga, M., Wever, K.E., 2015. A protocol format for the preparation, registration and publication of systematic reviews of animal intervention studies. Evidence-based Preclin. Med. 1, 1–9. http://dx.doi.org/10.1002/ebm2.7.

Demonstration of a study to Coordinate and Perform Human biomonitoring on a European Scale (DEMOCOPHES), 2013. Human Biomonitoring on a European Scale [WWW Document]. http://www.eu-hbm.info/euresult/media-corner/press-kit. (Accessed 18 April 2017).

Desvergne, B., Feige, J.N., Casals-Casas, C., 2009. PPAR-mediated activity of phthalates: a link to the obesity epidemic? Mol. Cell. Endocrinol. http://dx.doi.org/10.1016/j.mce.2009.02.017.

Dostal, L.A., Jenkins, W.L., Schwetz, B.A., 1987. Hepatic peroxisome proliferation and hypolipidemic effects of di(2-ethylhexyl)phthalate in neonatal and adult rats. Toxicol. Appl. Pharmacol. 87, 81–90.

Erythropel, H.C., Maric, M., Nicell, J.A., Leask, R.L., Yargeau, V., 2014. Leaching of the plasticizer di(2-ethylhexyl)phthalate (DEHP) from plastic containers and the question of human exposure. Appl. Microbiol. Biotechnol. http://dx.doi.org/10.1007/s00253-014-6183-8.

European Chemical Agency (ECHA), 2017. Dissemination Site - Substance Information: Bis(2-ethylhexyl) Phthalate [WWW Document]. https://echa.europa.eu/information-on-chemicals. (Accessed 10 July 2017).

European Chemical Agency (ECHA), 2014. Support Document for Identification of Bis(2-ethylhexyl) phthalate (DEHP) as a Substance of Very High Concern [WWW Document]. https://echa.europa.eu/documents/10162/fa429d23-21e7-4764-b223-6c8c98f8a01c. (Accessed 18 April 2017).

European Chemical Agency (ECHA), 2009. Data on manufacture, import, export, uses and releases of bis(2-ethylhexyl)phthalate (DEHP) as well as information on potential alternatives to its use [WWW Document]. https://echa.europa.eu/documents/10162/8fd5a74b-6807-42b6-ae1f-d1d7f04f40f8. (Accessed 18 April 2017).

European Chemical Agency (ECHA), 2008. Annex XV Dossier - Proposal for Identification of DEHP as an SVHC [WWW Document]. https://echa.europa.eu/documents/10162/e0a2db43-7a7f-4f18-92f9-2eaa3e4aff92. (Accessed 18 April 2017).

European Food Safety Authority (EFSA), 2005. 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. EFSA J. 243, 1–20.

Gore, A.C., Chappell, V.A., Fenton, S.E., Flaws, J.A., Nadal, A., Prins, G.S., Toppari, J., Zoeller, R.T., 2015. Executive summary to EDC-2: the endocrine Society's second scientific statement on endocrine-disrupting chemicals. Endocr. Rev. http://dx.doi.org/10.1210/er.2015-1093.

Grün, F., Blumberg, B., 2006. Environmental obesogens: organotins and endocrine disruption via nuclear receptor signaling. Endocrinology 147, S50–S55. http://dx.doi.org/10.1210/en.2005-1129.

Hao, C., Cheng, X., Guo, J., Xia, H., Ma, X., 2013. Perinatal exposure to diethyl-hexyl-phthalate induces obesity in mice. Front. Biosci. (Elite Ed.) 5, 725–733.

Hao, C., Cheng, X., Xia, H., Ma, X., 2012. The endocrine disruptor mono-(2-ethylhexyl) phthalate promotes adipocyte differentiation and induces obesity in mice. Biosci. Rep. 32, 619–629. http://dx.doi.org/10.1042/BSR20120042.

Heindel, J.J., Schug, T.T., 2013. The perfect storm for obesity. Obesity (Silver Spring) 21, 1079–1080.

Hooijmans, C.R., Leenaars, M., Ritskes-Hoitinga, M., 2010. A gold standard publication checklist to improve the quality of animal studies, to fully integrate the three Rs, and to make systematic reviews more feasible. ATLA Altern. Lab. Anim. 38, 167–182.

Hooijmans, C.R., Rovers, M.M., de Vries, R.B., Leenaars, M., Ritskes-Hoitinga, M., Langendam, M.W., 2014. SYRCLE's risk of bias tool for animal studies. BMC Med. Res. Methodol. 14, 43. http://dx.doi.org/10.1186/1471-2288-14-43.

Kilkenny, C., Browne, W.J., Cuthill, I.C., Emerson, M., Altman, D.G., 2010. Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research. PLoS Biol. 8, e1000412. http://dx.doi.org/10.1371/journal.pbio.1000412.

Kim, S.H., Park, M.J., 2014. Phthalate exposure and childhood obesity. Ann. Pediatr. Endocrinol. Metab. 19, 69–75. http://dx.doi.org/10.6065/apem.2014.19.2.69.

Li, L.X., Chen, L., Meng, X.Z., Chen, B.H., Chen, S.Q., Zhao, Y., Zhao, L.F., Liang, Y., Zhang, Y.H., 2013. Exposure levels of environmental endocrine disruptors in mother-newborn pairs in China and their placental transfer characteristics8. PLoS One 8. http://dx.doi.org/10.1371/journal.pone.0062526.

National Toxicology Program (NTP), 2015. Handbook for Conducting a Literature-based Health Assessment Using OHAT Approach for Systematic Review and Evidence Integration [WWW Document]. http://ntp.niehs.nih.gov/go/38673. (Accessed 9 January 2017).

National Toxicology Program (NTP), 2006. NTP-CERHR Monograph on the Potential Human Reproductive and Developmental Effects of Di(2-Ethylhexyl) Phthalate (DEHP). No. 06–4476.

US Environmental Protection Agency (US EPA), 1987. Integrated Risk Information System: Di(2-ethylhexyl)phthalate (DEHP) [WWW Document]. https://cfpub.epa.gov/ncea/iris/iris_documents/documents/subst/0014_summary.pdf. (Accessed 25 May 2017).

Wassenaar, P.N.H., Trasande, L., Legler, J., 2017. Systematic review and meta-analysis of early life exposure to bisphenol A and obesity related outcomes in rodents. Environ. Health Perspect. (in press).

Zota, A.R., Calafat, A.M., Woodruff, T.J., 2014. Temporal trends in phthalate exposures: findings from the national health and nutrition examination survey, 2001-2010. Environ. Health Perspect. 122, 235–241. http://dx.doi.org/10.1289/ehp.1306681.

Review Manager (RevMan), 2014. (Computer program). Version 5.3. Copenhagen: The Nordic Cochrane Centre. The Cochrane Collaboration.