

EPA Public Access

Author manuscript

Environ Int. Author manuscript; available in PMC 2021 December 01.

About author manuscripts

Submit a manuscript

Published in final edited form as:

Environ Int. 2020 December ; 145: 105848. doi:10.1016/j.envint.2020.105848.

Hazards of diethyl phthalate (DEP) exposure: A systematic review of animal toxicology studies

James A. Weaver^a, Brandiese E.J. Beverly^{a,1}, Nagalakshmi Keshava^a, Anuradha Mudipalli^a, Xabier Arzuaga^b, Christine Cai^b, Andrew K. Hotchkiss^a, Susan L. Makris^b, Erin E. Yost^{a,*} ^aU.S. Environmental Protection Agency, Center for Public Health and Environmental Assessment, Research Triangle Park, NC, United States

^bU.S. Environmental Protection Agency, Center for Public Health and Environmental Assessment, Washington, DC, United States

Abstract

Background: Diethyl phthalate (DEP) is widely used in many commercially available products including plastics and personal care products. DEP has generally not been found to share the antiandrogenic mode of action that is common among other types of phthalates, but there is emerging evidence that DEP may be associated with other types of health effects.

Objective: To inform chemical risk assessment, we performed a systematic review to identify and characterize outcomes within six broad hazard categories (male reproductive, female reproductive, developmental, liver, kidney, and cancer) following exposure of nonhuman mammalian animals to DEP or its primary metabolite, monoethyl phthalate (MEP).

Methods: A literature search was conducted in online scientific databases (PubMed, Web of Science, Toxline, Toxcenter) and Toxic Substances Control Act Submissions, augmented by review of online regulatory sources as well as forward and backward searches. Studies were selected for inclusion using PECO (Population, Exposure, Comparator, Outcome) criteria. Studies were evaluated using criteria defined a priori for reporting quality, risk of bias, and sensitivity using a domain-based approach. Evidence was synthesized by outcome and life stage of exposure, and strength of evidence was summarized into categories of *robust, moderate, slight, indeterminate*, or *compelling evidence of no effect*, using a structured framework.

*Corresponding author. Yost.Erin@epa.gov (E.E. Yost). .

Publisher's Disclaimer: 6. Disclaimer

Appendix A. Supplementary material

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/BY-NC-ND/4.0/).

¹Present affiliations: National Institute of Environmental Health Sciences, National Toxicology Program, Office of Health Assessment and Translation, Research Triangle Park, NC, United States.

The views expressed in this article are those of the author(s) and do not necessarily represent the views or policies of the U.S. Environmental Protection Agency.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envint.2020.105848.

Results: Thirty-four experimental studies in animals were included in this analysis. Although no effects on androgen-dependent male reproductive development were observed following gestational exposure to DEP, there was evidence including effects on sperm following peripubertal and adult exposures, and the overall evidence for male reproductive effects was considered *moderate*. There was *moderate* evidence that DEP exposure can lead to developmental effects, with the major effect being reduced postnatal growth following gestational or early postnatal exposure; this generally occurred at doses associated with maternal effects, consistent with the observation that DEP is not a potent developmental toxicant. The evidence for liver effects was considered *moderate* based on consistent changes in relative liver weight at higher dose levels; histopathological and biochemical changes indicative of hepatic effects were also observed, but primarily in studies that had significant concerns for risk of bias and sensitivity. The evidence for female reproductive effects was considered *slight* based on few reports of statistically significant effects on maternal body weight gain, organ weight changes, and pregnancy outcomes. Evidence for cancer and effects on kidney were judged to be *indeterminate* based on limited evidence (i.e., a single two-year cancer bioassay) and inconsistent findings, respectively.

Conclusions: These results suggest that DEP exposure may induce androgen-independent male reproductive toxicity (i.e., sperm effects) as well as developmental toxicity and hepatic effects, with some evidence of female reproductive toxicity. More research is warranted to fully evaluate these outcomes and strengthen confidence in this database.

Keywords

Phthalates; Diethyl phthalate; Systematic review; Risk assessment

1. Introduction

Diethyl phthalate (DEP), a colorless, odorless oily substance, is used to improve the performance and durability of many products (Consumer Product Safety Commission 2011; World Health Organization 2003). As a plasticizer, it is added to plastic polymers to help maintain flexibility. It has been used in a variety of products including plastic films, rubber, tape, toothbrushes, automotive components, tool handles and toys. In addition to plastics, DEP is present in a wide range of personal care products (e.g., cosmetics, perfume, hair spray, nail polish, soap, detergent, and lotions), industrial materials (e.g., rocket propellant, dyes, packaging, sealants and lubricants), and medical products (e.g., enteric coatings on tablets and in dental impression materials).

Phthalates including DEP are not covalently bound to products, and therefore are readily released into the environment where they may be absorbed orally, by inhalation, or dermally (Clark et al. 2011; Wormuth et al. 2006). Following oral exposure in rats, DEP primarily locates to the kidneys and liver followed by deposition in fat (Singh et al. 1975). DEP is rapidly metabolized into the active metabolite monoethyl phthalate (MEP), which is ultimately excreted into the urine and serves as a biomarker of DEP exposure. Exposure assessment data from the National Health and Nutrition Examination Survey (NHANES) show that MEP was detected in the urine of at least 98% of participants in the US general population in each survey cycle between 2001 and 2010; urinary concentrations of MEP declined significantly over that time, with a more pronounced trend towards decreased

urinary MEP in adults and adolescents compared to children, perhaps reflecting a trend towards decreased use of DEP in personal care products (Zota et al. 2014). Despite this trend, MEP levels tended to remain higher compared to other phthalate metabolites in urine across age groups in this study (Zota et al. 2014). Other analyses of biomonitoring data have similarly found that DEP is among the highest phthalate exposures for women of childbearing age, with personal care products being a major source (Consumer Product Safety Commission 2014; National Research Council 2008).

Unlike multiple other phthalates, DEP has not been found to inhibit fetal testosterone production (Consumer Product Safety Commission 2014), which is one of the major mechanisms underlying the "phthalate syndrome" phenotype that is observed in male rats following gestational exposure to certain phthalates. Phthalate syndrome is characterized by a spectrum of effects including underdevelopment of male reproductive organs, decreased anogenital distance (AGD), female-like nipple retention, cryptorchidism, and germ cell toxicity (Consumer Product Safety Commission 2014; National Research Council 2008). These effects are driven not only by a phthalate-induced decrease in testicular testosterone production, but also by decreased production of insulin-like-3 hormone and disrupted seminiferous cord formation, Sertoli cells, and germ cell development, which occur independently of changes in androgen production (Johnson et al. 2012; Martino-Andrade and Chahoud 2010; National Research Council 2008). A recent review by a Chronic Hazard Advisory Panel (CHAP), which focused on phthalates and phthalate alternatives used in children's toys and child care products, found that DEP does not cause phthalate syndrome in rats, although decreased testosterone and effects on sperm were observed in some studies in adult and peripubertal animals, and there were some associations between urinary MEP and male reproductive outcomes in humans (Consumer Product Safety Commission 2014). Based on these findings, the CHAP did not recommend that DEP be banned from children's toys and childcare products, particularly because these products are considered a negligible source of DEP exposure. However, the CHAP concluded that since exposures from personal care products, diet, and some pharmaceuticals can be substantial, exposure to DEP remains a concern and warrants further evaluation (Consumer Product Safety Commission 2014).

The aim of this study is to characterize the range of health outcomes associated with DEP exposure in animal toxicology studies using systematic review methods, which were not used in the CHAP review of phthalates. Our analysis includes evaluation of effects in both males and females and for all life stages of exposure, focusing on six broad hazard categories: male reproductive, female reproductive, developmental, liver, kidney, cancer. These hazard categories have been identified by our group as being broadly associated with phthalate exposure and were selected a priori due to the likelihood that they would have sufficient data available for hazard identification. The results can inform chemical risk assessment by providing a more comprehensive understanding of the effects of DEP exposure and identifying gaps in the currently available literature.

2. Methods

Our group has published several systematic reviews of epidemiological and animal toxicological studies describing health effects following exposure to phthalates (Radke et

al. 2018; Radke et al. 2019a, 2019b; Radke et al. 2020; Yost et al., 2019). This systematic review continues the evaluation of health effects after exposure to phthalates by focusing on animal studies of DEP. The literature searches and screening, study evaluation, data extraction, and evidence synthesis methods are described in detail in the systematic review protocol (provided as a supplementary file) and summarized here. The systematic review protocol also provides detailed definitions for the terminology used to describe study evaluation and evidence synthesis, which are summarized in Fig. 1. For easier reference, these definitions and key methods from the protocol related to study evaluation and evidence synthesis are also summarized in a separate supplementary file ("key methods supplement"). The systematic review methodology used here has been reviewed previously by the National Academies of Sciences, Engineering, and Medicine (2018) and was used in the systematic review of animal studies for diisobutyl phthalate (Yost et al. 2019), and reporting is in accordance with the checklist from PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) (http://www.prisma-statement.org/).

2.1. Literature searches and screening

A literature search was conducted in online scientific databases [PubMed, Web of Science, Toxline, Toxcenter] and Toxic Substances Control Act Test Submissions (TSCATS),² using search terms designed to capture all potentially pertinent studies. Initial database searches were conducted as early as March 2012, with updates of PubMed, Web of Science, and Toxline performed periodically through January 2020 (see protocol Section 3). The literature search strings used for PubMed and Web of Science evolved over time, but the final literature update in January 2020 was conducted using additional chemical synonyms and with topical terms removed and with no date limitations in order to capture a broader array of studies. The results of this literature search were supplemented by forward and backward searches, searching citations from key references, manual search of citations from key regulatory documents, and by addition of references that had been previously identified from an earlier DEP review effort and added to EPA's Health and Environmental Research Online (HERO) database (https://hero.epa.gov/hero/index.cfm/project/page/project_id/1097). The number of studies identified from each source and literature update can be found on the HERO project page.

A PECO (Population, Exposure, Comparator, Outcome) was developed to frame the research question and guide the screening of relevant studies. The PECO identifies the following as the inclusion criteria for the systematic review of DEP animal toxicology studies (see protocol Section 2 for the full PECO):

- <u>Population</u>: Nonhuman mammalian animal species (whole organism) of any life stage.
- <u>Exposure</u>: Any administered dose of DEP or MEP as singular compounds, via oral, dermal, or inhalation routes of exposure.
- <u>Comparator</u>: Exposure to vehicle-only or untreated control

 $^{^{2}}$ TSCATS records were identified by searching EPA's TSCATS2 database, searching TSCATS1 via Toxline, and conducting Google searches for TSCA section 8e recent notices.

Environ Int. Author manuscript; available in PMC 2021 December 01.

• <u>Outcome</u>: Any examination of male reproductive, female reproductive, developmental, liver, kidney, or cancer outcomes.

Title/abstract and full text screening were performed by two reviewers, and all identified animal toxicology studies underwent full-text screening to determine compliance with the PECO. Peer-reviewed studies that contained original data and complied with the PECO were selected for inclusion and were moved forward for study evaluation. Studies providing supporting health effects data (e.g. mechanistic, genotoxic, or toxicokinetic studies) were also compiled in HERO and annotated during the screening process.

2.2. Study evaluation

For each study selected for inclusion, the quality and informativeness of the evidence was rated by evaluating domains related to reporting quality, risk of bias, and sensitivity (see protocol Section 4; abbreviated version available in the key methods supplement). Evaluations first considered reporting quality, which refers to whether the study has reported sufficient details to conduct a risk of bias and sensitivity analysis; if a study does not report critical information (e.g. species, test article name) it may be excluded from further consideration. Risk of bias, sometimes referred to as internal validity, is the extent to which the design or conduct of a study may alter the ability to provide accurate (unbiased) evidence to support the relationship between exposure and effects (Higgins 2011). Sensitivity refers to the extent to which a study is likely to detect a true effect caused by exposure (Cooper et al. 2016; Higgins 2011).

All study evaluation ratings are documented and publicly available in EPA's version of Health Assessment Workspace Collaborative (HAWC), a free and open source webbased software application (https://hawcprd.epa.gov/assessment/552/). Study evaluation was conducted in the following domains: reporting quality; allocation; observational bias/blinding; confounding; selective reporting and attrition; chemical administration and characterization; exposure timing, frequency and duration; endpoint sensitivity and specificity; and results presentation. For each domain, core questions and basic considerations provided guidance on how a reviewer might evaluate and judge a study for that domain (see Table 9 of the protocol or Table A of the key methods supplement).

At least two reviewers independently assessed each study, and any conflicts were resolved through discussion among reviewers or other technical experts. When information needed for the evaluation was missing from a key study, an attempt was made to contact the study authors for clarification. All communication with study authors was documented and is available in HERO (tagged as Personal Correspondence with Authors) and was annotated in HAWC whenever it was used to inform a study evaluation.

For each study, in each evaluation domain, reviewers reached a consensus on a rating of *Good, Adequate, Deficient*, or *Critically Deficient*. These individual ratings were then combined to reach an overall study confidence classification of *High, Medium, Low*, or *Uninformative*. The evaluation process was performed separately for each outcome reported in a study, as the utility of a study may vary for different outcomes.

2.3. Data extraction

Data from included studies were extracted into HAWC (https://hawcprd.epa.gov/ assessment/552/). Dose levels are presented as mg/kg-day. For dietary exposure studies that reported dose levels as concentrations, dose conversions to mg/kg-day were made using US EPA default food or water consumption rates and body weights for the species/strain and sex of the animal of interest (US EPA 1988).

2.4. Evidence synthesis

For each outcome, the available evidence from the included animal studies was synthesized using a narrative approach, using the following considerations to articulate the strengths and weaknesses of the available evidence [adapted from (Hill 1965)]: consistency, biological gradient (dose–response), strength (effect magnitude) and precision, biological plausibility, and coherence. These considerations are defined in Table 10 of the protocol, and Table 11 of the protocol provides more information about how they were used to characterize the strength of evidence. When possible, an effort was made to evaluate the data according to the age and developmental stage of exposure to account for life stage-specific windows of susceptibility, as recommended by EPA's Framework for Assessing Health Risk of Environmental Exposures to Children (US EPA 2006) and by Makris et al. (2008). When available, informative mechanistic data were used to augment the qualitative syntheses.

Based on this synthesis, each outcome was assigned a strength of evidence conclusion of Robust, Moderate, Slight, Indeterminate, or Compelling evidence of no effect (see definitions of these terms in Table 13 of the protocol or Table B of the key methods supplement). Robust and Moderate describe evidence that supports a hazard, differentiated by the quantity and quality of information available to rule out alternative explanations for the results (Robust describes an evidence base for which there is reasonable confidence that results are not due to chance, bias, or confounding; whereas Moderate describes an evidence base that has greater uncertainty, e.g. due to a smaller number of studies or some heterogeneity in results). Slight evidence includes situations in which there is some evidence that supports a hazard, but a conclusion of *Moderate* does not apply. *Indeterminate* describes a situation where the evidence is limited or inconsistent and cannot provide a basis for making a conclusion in either direction, or if the available studies are largely null but do not reach the level required to conclude there is compelling evidence of no effect. Compelling evidence of no effect represents a situation where extensive evidence across a range of populations and exposures identified no association between exposure and hazard. The ratings for individual outcomes were then summarized into an overall strength of evidence conclusion for each of the six hazards (male reproductive, female reproductive, developmental, liver, kidney, cancer). Rationales for strength of evidence conclusions are presented in evidence profile tables using a structured format based on the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach for evaluating certainty in the evidence (Guyatt et al. 2011; Schünemann et al. 2011).

3. Results

3.1. Study selection

Literature search and screening results are summarized in Fig. 2. The literature search retrieved a total of 3,052 unique records for DEP, of which 79 were identified as animal health effect studies. Of the animal health effect studies, 39 were excluded for not meeting PECO criteria, e.g. skin and eye irritation studies and studies that used non-mammalian species. Three sets of the included studies were found to be multiple publications of the same data [Hazleton Laboratories America Inc (1983a, 1983b, 1992) and Hardin et al. (1987); Lamb et al. (1987) and RTI International (1984); and Field et al. (1993) and NTP (1988)] and were considered thereafter to be a single study [cited here as Hardin et al. (1987), RTI International (1984), and NTP (1988)]. At the study evaluation phase, one study was found to be uninformative due to critical deficiencies in the presentation of results and was excluded from further analysis (Hu et al. 2018). Therefore, a total of 34 studies were included this analysis (Table 1).

3.2. Summary of included studies

The included studies are summarized in Table 1. The database of DEP studies is diverse and consists of multigenerational studies, gestational exposure studies, and studies in peripubertal or adult animals. All studies were either oral or dermal exposures conducted in rats, rabbits or mice. Kwack et al. (2009) and Kwack et al. (2010) exposed animals to either DEP or MEP, and all other studies used DEP.

Eight studies assessed multigenerational effects of DEP exposure. Of these, reproductive and developmental effects were evaluated in the continuous breeding study in mice by RTI International (1984) and in the two-generation reproduction study in rats by Fujii et al. (2005), which exposed animals beginning prior to the mating of F0 parental animals and continuing through the weaning of F2 offspring. Four multigenerational studies report data from the same group of animals exposed to a low dose of DEP (approximately 2.85 mg/kg-day) from the F0 to F2 generations, focusing on either hepatic toxicity (Pereira et al., 2007a, 2007c; Pereira and Rao, 2007) or reproductive toxicity (Pereira et al., 2007b). The remaining two multigenerational studies (Hu et al. 2016; Manservisi et al. 2015) exposed F1 female rats to a low dose (0.173 mg/kg-day) of DEP beginning at postnatal day (PND) 1 via lactation, and continuing through PND 181 via oral gavage; Manservisi et al. (2015) reported on mammary gland effects and fertility in F1 females and growth and survival of F2 offspring, and Hu et al. (2016) reported body weight data for a subset of these same F1 female rats at PNDs 62 and 181.

Effects in developing animals were also evaluated in studies that exposed animals during gestation and/or early postnatal life. Of these, the studies by Gray et al. (2000), Howdeshell et al. (2008), Furr et al. (2014), and Liu et al. (2005) exposed rats during late gestation [gestation day (GD) 14 – PND 3, GD 8–18, GD 14 – 18, and GD 12–19, respectively] and focused on male reproductive development and phthalate syndrome effects. These exposures coincide with the critical window of male sexual differentiation (~GD 14–18), which is known to be the sensitive window of exposure for the induction of phthalate syndrome.

Hardin et al. (1987) evaluated fetal survival and growth in mice exposed to a single high dose of DEP from gestation day (GD) 6–13. NTP (1988) and Procter & Gamble (1994) evaluated fetal survival, growth, and structural alterations in rats and rabbits following exposure from GD 6–15 and 6–18, respectively. Setti Ahmed et al. (2018) evaluated intestinal morphology and organ weight changes in rats exposed from GD 8 – PND 30. Most of these developmental exposure studies also provided relevant data on maternal reproductive endpoints (e.g. maternal body weight gain, pregnancy outcomes) in addition to effects on the developing animals.

The remaining studies evaluated rats or mice following peripubertal or adult exposure. Oishi and Hiraga (1980), Mondal et al. (2019), and Jones et al. (1993) focused on male reproductive effects (testosterone, sperm parameters, and/or testicular histology), and Shiraishi et al. (2006) evaluated reproductive effects in both male and female rats including hormone levels, sperm parameters, and estrous cyclicity. Kwack et al. (2010) and Kwack et al. (2009) evaluated general toxicity (organ weights, serum biochemistry, urinalysis) in DEPand MEP-exposed rats, with a sperm evaluation also conducted in Kwack et al. (2009). A two-year dermal exposure study in rats and mice by NTP (1995) provided information on tumor incidence. Other studies focused on general toxicity and hepatic effects, including organ weight, histopathology, and biochemical changes. This includes a series of low dose studies in rats or mice performed by the same laboratory (Mapuskar et al., 2007; Pereira et al., 2006, 2008a, 2008b, Pereira and Rao, 2006a, 2006b, 2007; Sinkar and Rao, 2007; Sonde et al., 2000), as well as the studies in rats by Moody and Reddy (1978), Moody and Reddy (1982), and Brown et al. (1978).

3.3. Study evaluation

Overall study confidence classifications by outcome are summarized in Table 1, and heat maps summarizing study evaluation ratings by domain are provided in the Supplementary Materials (Fig. S1). Fig. S1 provides links to an interactive figure in HAWC, where rationale for the study evaluation ratings is documented.

Based on the study evaluation considerations outlined in the systematic review protocol, confidence was reduced in some studies that had incomplete reporting of experimental designs or results. For example, dose-related histopathological findings were frequently reported qualitatively, with no quantitative data provided on the incidence or severity of lesions; such outcomes were considered *low confidence*. Additionally, some study outcomes were rated *medium* or *low confidence* due to sensitivity concerns with certain endpoint measurements. For instance, for evaluation of maternal body weight gain, rabbits were not considered an appropriate test species since body weight changes in rabbits are more variable compared to other species (US EPA 1991); and, for all species, confidence was reduced in the maternal body weight gain measurements in studies that did not adjust for gravid uterine weight, which facilitates the interpretation of maternal toxicity relative to effects on fetal body weight (US EPA 1991). For organ weights, it was considered best practice for studies to report both absolute and relative (adjusted to body weight) measurements, although relative organ weight as a standalone measurement was generally considered to be acceptable for most organs (Bailey et al. 2004). For testis, however, studies

only reporting relative organ weight were considered *low confidence*, since testis weight is not proportionate to body weight (Bailey et al. 2004).

Specific concerns were raised regarding studies from two laboratory groups that tested relatively low doses of DEP; these include concerns about exposure characterization as well as other concerns for risk of bias and sensitivity. The studies by Manservisi et al. (2015) and Hu et al. (2016), both of which reported data on the same group of animals, dosed rats with 0.173 mg/kg-day three times per week from birth through PND 181 and stated via personal correspondence that they verified the concentration of DEP in the dosing solutions but did not evaluate background levels of DEP that may be present due to the use of phthalate in plastics or other environmental sources. This is potentially problematic because a separate dose range-finding study by these authors (Teitelbaum et al. 2016) reported elevated levels of MEP in the urine of control animals, which suggests the potential for background exposures or contamination that could significantly impact the nominal dose levels. A series of studies by another laboratory exposed rats or mice to low doses of DEP in diet (0.57–6.25 mg/kg-day) (Mapuskar et al., 2007; Pereira et al., 2006, 2007a, 2007b, 2007c, 2008a, 2008b, Pereira and Rao, 2006a, 2006b, 2007) or to 50 ppm DEP in drinking water (Sinkar and Rao 2007; Sonde et al. 2000) without verifying the nominal doses. Again, concerns were raised that undetected background levels of phthalates in control groups could mask true effects and reduce study sensitivity. Moreover, several of these studies by Pereira and coauthors stated that the concentration of DEP in the diet was increased on a weekly basis to maintain the dose in proportion with the animals' body weight but did not provide additional information on this dose adjustment, which raises separate questions about the accuracy of the nominal doses in these studies. Studies by these two laboratory groups also had other concerns raised during study evaluation; for instance, data from littermates were presented as the average of individual pups, which has the potential to overestimate the statistical significance of experimental findings (Haseman et al. 2001). Most of these studies also described histopathological changes in the exposed animals without providing quantitative data to support their findings. Overall, due to these cumulative concerns, these studies were rated as low confidence for all reported outcomes.

3.4. Male reproductive effects

Male reproductive effects were evaluated according to the life stage of exposure (US EPA 2006). Figs. indicating the doses at which statistically significant male reproductive effects occurred are provided in Supplemental Materials (Figs. S2–S4).

3.4.1. Summary of gestational and early postnatal exposure studies (including F1 or F2 offspring from multigenerational studies)—Three *high* or *medium confidence* studies investigated effects on testosterone production or levels in male rats following gestational exposure (Furr et al. 2014; Howdeshell et al. 2008) or exposure from gestation through PND 3 (Gray et al. 2000). These studies all used exposure periods that included the critical window of male sexual differentiation in late gestation (which occurs between ~ GD 14–18), and therefore are considered relevant for the evaluation of fetal testosterone production and other phthalate syndrome effects. No statistically significant effects on fetal testicular testosterone production were observed at GD 18

following exposure from GD 14-18 (Furr et al. 2014) or GD 8-18 (Howdeshell et al. 2008) at doses up to 750 and 900 mg/kg-day, respectively. Likewise, there was no statistically significant effect on serum testosterone levels in adult male rats that had been exposed to 750 mg/kg-day DEP from GD 14 - PND 3 (Gray et al. 2000). Additional supporting mechanistic evidence was provided by Liu et al. (2005), who did not report testosterone production but evaluated global gene expression in the fetal rat testis after exposure to 500 mg/kg-day from GD 12–19; gene expression in DEP-treated animals was similar to controls, whereas known antiandrogenic phthalates (e.g. dibutyl phthalate, diethylhexyl phthalate) downregulated expression of genes involved in cholesterol homeostasis and steroidogenesis (Liu et al. 2005). A follow-up study by Clewell et al. (2010) evaluated testicular MEP levels in a subset of male fetuses from the study by Liu et al. (2005) and found MEP present at relatively high levels, indicating that the lack of effect of DEP in the study by Liu et al. (2005) was not due to the dose not reaching the testis. Taken together, these studies consistently suggest lack of effect on fetal testosterone production. Rats have been found to be more sensitive to the anti-androgenic effects of phthalates compared to mice (Johnson et al. 2012), so the lack of effects in rats strengthens the evidence that DEP does not affect fetal testosterone production; however, additional evidence (e.g. studies in other species besides rat) would be needed to have confidence in a conclusion of *compelling evidence* of no effect. The evidence for effects on testosterone levels after gestational exposure was therefore considered indeterminate.

Male reproductive organ weights and other biomarkers of androgen-dependent reproductive development were evaluated in F1 peripubertal and adult rats after exposure from GD 14 – PND 3 (Gray et al. 2000) and in F1 and F2 weanling rats in the two-generation reproduction study (Fujii et al. 2005). Additionally, AGD was evaluated in fetal rats after exposure from GD 12-19 (Liu et al. 2005). All of these studies were judged to be high confidence for these outcomes. There were no statistically significant effects on male reproductive organ weights (Fujii et al. 2005; Gray et al. 2000), with the exception of a 20% decrease in absolute prostate weights and 17% increase in relative seminal vesical weight observed in F1 weanlings exposed to 1016 mg/kg-day DEP (Fujii et al. 2005). Prostate and seminal vesicle weights were not affected in F2 weanlings from the same study (Fujii et al. 2005). There were no effects on AGD (Fujii et al. 2005; Gray et al. 2000; Liu et al. 2005) or nipple retention (Gray et al. 2000), both of which are biomarkers of androgen-dependent reproductive development. Gestational exposure to DEP did not affect the timing of sexual maturation, as measured by the age at onset of preputial separation (Fujii et al. 2005; Gray et al. 2000). The lack of effect on these outcomes is consistent with the observation that DEP does not affect testosterone in rats exposed during gestation. However, since a larger body of evidence would be needed to conclude there was compelling evidence of no effect, the evidence for effects on male reproductive organ weights and morphological development were both considered indeterminate.

3.4.2. Summary of peripubertal and adult exposure studies (including F0 or F1 parental animals from multigenerational studies)—In contrast with results observed after gestational exposure, several studies reported decreased testosterone levels following peripubertal or adult exposure to DEP in male rats. In a *high confidence* two-

generation reproduction study, Fujii et al. (2005) observed decreased serum testosterone in F0 males following 14 weeks of exposure in all dose groups, reaching statistical significance in the two highest dose groups (197 and 1016 mg/kg-day); however, this evaluation was performed on a relatively small subset of the animals (n = 6/group) and results showed a large amount of variability with a nonmonotonic magnitude of change (decreased by 80% in the 197 mg/kg-day group and 50% in the 1016 mg/kg-day group). This study did not evaluate hormone levels in F1 animals, so it is not clear whether this effect was present across other generations; however, it is notable that there was minimal or no effect in these animals on male reproductive organ weights (discussed below), which are known to be sensitive to changes in androgen levels (US EPA 1996). Dose-related statistically significant decreases in serum testosterone and androstenedione were also observed in adult male rats exposed to 0.57 mg/kg-day to 2.85 mg/kg-day DEP for 150 days in the study by Pereira et al. (2008a), which is considered low confidence for this outcome due to the study design concerns described in Section 3.3. In addition, serum and testicular testosterone levels were statistically significantly decreased in young rats exposed to DEP for 7 days (Oishi and Hiraga 1980), although this data was presented as a percentage of control without a measure of variance and is considered low confidence. In contrast to other findings, no effect on serum testosterone or gonadotropin levels was observed in a *high confidence* study in Wistar rats following exposures up to 1000 mg/kg-day DEP for 28 days (Shiraishi et al. 2006), although those authors did observe a statistically significant decrease in serum estradiol for males in the 1000 mg/kg-day dose group. No effects on serum testosterone were observed in a *medium* confidence study in Swiss albino mice following exposures up to 10 mg/kg-day in diet for 3 months (Mondal et al. 2019). Additionally, mechanistic evidence to support a lack of effect on testicular steroidogenesis was available in the study in rats by Foster et al. (1983), which found that peripubertal exposure to $\sim 1600 \text{ mg/kg-day DEP}$ for up to 4 days did not affect the activity of testicular steroidogenic enzymes, whereas exposure to a known antiandrogenic phthalate (dipentyl phthalate) decreased the enzyme activity in a time-dependent manner; and Jones et al. (1993) reported that MEP did not decrease testosterone production in an in vitro culture of rat Levdig cells. Overall, although decreased testosterone was observed in three studies, two of these studies had significant concerns for bias and the results are not supported by mechanistic evidence or other coherent effects (e.g. effects on testis weight) in the high confidence study by Fujii et al. (2005). The evidence for effects on testosterone after peripubertal or adult exposure was considered *slight*.

Four *high* or *medium confidence* studies that evaluated sperm parameters in rats or mice found statistically significant effects on sperm count, motility, or morphology following multigenerational exposure to DEP (Fujii et al. 2005; RTI International 1984), peripubertal exposure to DEP or MEP (Kwack et al. 2009), or adult exposure to DEP (Mondal et al. 2019). The *high confidence* study by Fujii et al. (2005) observed a statistically significant increase in abnormal or tailless sperm in both F0 and F1 rats, although the magnitude of effect was low (abnormal sperm rate was 1.52% in F1 animals in the 1150 mg/kg-day group versus 0.6% in control animals) and effects in F0 animals were nonmonotonic (occurring at 197 mg/kg-day group but not 1016 mg/kg-day). Sperm counts and motility were not affected in this study. The *high confidence* study by Kwack et al. (2009) observed that epididymal sperm count and percent motility were statistically significantly decreased in

rats following exposure to 250 mg/kg-day MEP for four weeks, while percent linearity was statistically significantly decreased at 500 mg/kg-day DEP, with no effects on other sperm parameters. The *high confidence* study by RTI International (1984) reported that epididymal sperm counts were statistically significantly decreased in F1 mice dosed with 3640 mg/kg-day, with no effects on percent motility, abnormal sperm, or tailless sperm. The *medium confidence* study by Mondal et al. (2019) reported a statistically significant decrease in epididymal sperm count, motility, and viability and increased sperm morphological abnormalities in adult mice dosed with 1 or 10 mg/kg-day for three months, and provided supporting mechanistic evidence demonstrating oxidative stress in germ cells and apoptosis in testicular sections. Conversely, the *medium confidence* study by Shiraishi et al. (2006) reported no effects on epididymal sperm counts or morphology in adult rats following 28-day DEP exposure at doses up to 1000 mg/kg-day, although the authors provided no quantitative data. Although effects varied across studies, alterations in sperm quality were observed across most studies including three considered *high confidence*, and therefore the evidence for effects on sperm parameters is considered *moderate*.

No effects on copulation or fertility in F0 or F1 mating pairs were observed in two multigenerational studies that are considered *high confidence* for this endpoint (Fujii et al. 2005; RTI International 1984). Since only two studies were available, the evidence for effects on male fertility is considered *indeterminate*.

Reproductive organ weights in males following adult exposure to DEP were measured in several studies. In high confidence studies that reported absolute organ weights, DEP generally did not affect testosterone-dependent male reproductive organ weights (e.g., testes, epididymides, prostate, seminal vesicles) in F0 and F1 parental rats and mice (Fujii et al. 2005; RTI International 1984), although Fujii et al. (2005) reported a slight (5%) but statistically significant decrease in epididymal weights in F0 males at 1016 mg/kg-day. The medium confidence study by Oishi and Hiraga (1980) found no effect on absolute or relative testis weight in young rats exposed to DEP for 7 days. In low confidence studies, effects were inconsistent. The multigenerational study in rats by Pereira et al. (2007d) reported statistically significantly decreased absolute testis weights in F0 parental males but not adult F1 males exposed to 2.85 mg/kg-day in diet, and Pereira et al. (2008a) reported a statistically significant dose-related decrease in absolute testis and epididymis weights in adult male rats after exposure to 0.57 to 2.85 mg/kg-day. Brown et al. (1978) reported a statistically significant increase in relative testis weights in rats after 2, 6, or 16 weeks of exposure to 3160 mg/kg-day; however, it is possible that this is an artifact of decreased body weight in these animals, since testis and body weights are not proportional (Bailey et al. 2004). In other studies that did not observe significant effects on body weight, relative testis weights were not affected by DEP (Kwack et al. 2009; Kwack et al. 2010; Shiraishi et al. 2006). Given that absolute organ weights were largely unaffected in *high* confidence studies in rats and mice, the evidence for effects on male reproductive organ weights is considered indeterminate.

In addition, five studies performed histopathological evaluations of male reproductive organs. The *low confidence study* by Mondal et al. (2019) reported altered testicular histoarchitecture (azoospermia, loss of tubular compactness and increased vacuolization,

thinning of basement membrane), which is coherent with effects on epididymal sperm in this study. The *medium confidence* study by Jones et al. (1993) reported no changes in seminiferous tubular structure or Leydig cell morphology in rats exposed to 2000 mg/ kg-day for two days, although there were effects on Leydig cell cytoplasmic ultrastructure (mitochondrial swelling, focal dilation and vesiculation of the smooth endoplasmic reticulum). The three remaining *medium* or *low confidence* studies did not observe any treatment-related effects (Brown et al. 1978; Fujii et al. 2005; Shiraishi et al. 2006). None of these studies provided quantitative data to support their observations. Given this limited and inconsistent dataset, the evidence for histopathological effects in male reproductive organs is considered *indeterminant*.

3.4.3. Synthesis of results for male reproductive effects—Overall, the available studies suggest that there is *moderate* evidence that DEP is a male reproductive toxicant (Table 2). Decreased sperm quantity or quality was observed across almost all studies that evaluated sperm parameters, although the magnitude of effect was low in some cases and there was variation in the parameters affected across studies. The observed effects on sperm are nevertheless compelling because similar effects have been reported for other phthalates, and are thought to result from direct targeting of Sertoli cells that occurs via an androgen-independent mode of action (Johnson et al. 2012; National Research Council 2008). It is plausible that DEP may operate through a similar mechanism, leading to effects on sperm in absence of an effect on steroidogenesis. It is also possible that this reproductive toxicity is mediated by an androgen-dependent mechanism, given that decreased testosterone was also observed in three peripubertal or adult exposure studies; however, there is less confidence in the finding that DEP exposure may affect hormone levels. Gestational exposure studies did not show effects of DEP on testosterone production, biomarkers of male reproductive development, sexual maturation, or (to any great extent) male reproductive organ weights.

Despite the effects on sperm parameters, there is no evidence that fertility was affected in the two-generation study by Fujii et al. (2005) or the continuous breeding study by RTI International (1984). However, it has been demonstrated that rodents can remain fertile even after dramatic reductions in sperm counts, whereas a relatively small change in sperm count may impact human fertility (Mangelsdorf et al. 2003; Sharpe 2010). Therefore, the findings of decreased sperm quality in animal models is of potential relevance for human health risk assessment.

Strengths of this evidence base include the availability of several *high confidence* gestational exposure studies, with two large multigenerational studies that assessed multiple outcomes in postnatal and adult animals (Fujii et al. 2005; RTI International 1984). Other strengths are the availability of studies in both rats and mice that assessed similar outcomes, although the majority of evidence was from studies in rats.

3.5. Female reproductive effects

Figs. indicating the doses at which statistically significant female reproductive effects occurred are provided in the Supplementary Materials (Figs. S5–S7 and S18).

3.5.1 . Summary of available studies—Effects on pregnancy outcomes (including mating, fertility, fecundity, and gestation length) were evaluated in two high confidence studies following continuous DEP exposure across multiple generations in rats (Fujii et al. 2005) and mice (RTI International 1984), and in one low confidence study that evaluated effects in F1 female rats that had been exposed to a low dose of DEP (0.173 mg/kg-day) since birth (Manservisi et al. 2015). In the two-generation study in rats, Fujii et al. (2005) observed that F1 parental females had a slight but statistically significant decrease in gestation length after exposure to 1375 mg/kg-day DEP. This effect was not observed in F0 parental females, and there were no effects on copulation, fertility, or litter size in either generation. In the continuous breeding study in mice (RTI International 1984), litter size was statistically significantly reduced by 14% in F1 parental females following exposure to 3640 mg/kg-day. There were no effects on litter size or number of litters per mating pair for F0 parental females, and no effect on fertility (number fertile per number cohabited) in either F0 or F1 animals. Manservisi et al. (2015) found that pregnancy rate was not affected, and that litter size was statistically significantly increased in DEP-treated females compared to controls, although the finding for litter size is considered low confidence due to the concerns described in Section 3.3. Given that some effects on litter size and gestation length in F1 parental females were observed in two high confidence studies, the evidence for effects on pregnancy outcomes was considered slight.

Maternal body weight parameters were assessed in the two-generation study in rats by Fujii et al. (2005) and in several studies that exposed animals during gestation only. In rats exposed to DEP from GD 6–15, NTP (1988) reported a decreasing trend in maternal body weight gain after correcting for gravid uterine weight, indicating that the effect was maternal rather than fetal. In contrast, the multigenerational exposure study by Fujii et al. (2005) reported a statistically significant increase in F0 maternal body weight gain in several dose groups, but no effects on maternal weight gain in the F1 generation. No effects on maternal weight gain were observed in the remaining studies in mice (Hardin et al. 1987), rats (Furr et al. 2014; Gray et al. 2000; Howdeshell et al. 2008), or rabbits (Procter & Gamble 1994), although the latter study should be interpreted with additional caution, since maternal body weight during pregnancy can be highly variable in rabbits (US EPA 1991). Taken together, DEP may affect maternal weight gain at high doses in the rat, but effects were inconsistent, possibly due to interspecies sensitivity differences or differences in experimental design. Therefore, the evidence for effects on maternal body weight gain are considered *indeterminate*.

Multiple studies evaluated organ weights in females that had been exposed to DEP during development or as adults. No DEP-related effects were reported for gravid uterine weights in adult pregnant rats or rabbits (NTP 1988; Procter & Gamble 1994) or absolute or relative uterine or ovarian weights in adult rats (Brown et al. 1978; Shiraishi et al. 2006), F0 or F1 adult rats (Fujii et al. 2005), or F1 adult mice (RTI International 1984). In contrast, absolute uterine weights were statistically significantly reduced in F1 and F2 weanling female rats after gestational exposure to 1297 mg/kg-day and 1375 mg/kg-day, respectively, although this effect appears to be transient, since effects on uterine weight were not observed in adult F1 females at necropsy (Fujii et al. 2005). Statistically significant increases in relative

ovary weights in F0 and F1 female rats following exposure to 2.85 mg/kg-day and 1.425 mg/kg-day, respectively, were reported in the multigenerational study by Pereira et al. (2007d), although this study is considered *low confidence* due to the concerns discussed in Section 3.3. Taken together, while there is some evidence for effects on various organ weights from both *high* and *low* confidence studies, most studies do not support that DEP induces permanent, adverse effects. The evidence for effects on female reproductive organ weight is considered *slight*.

In a histopathological evaluation, the *low confidence* study by Manservisi et al. (2015) reported mammary gland effects (decreases in the size of lobular structures and a darker and denser appearance of these structures due to the lower dilation of the secretory alveoli) in parous F1 females that had been exposed to 0.173 mg/kg-day DEP since birth, although the sample size was small (n = 3/group) and only semi-quantitative results were presented; corresponding effects in nulliparous females after lactational and direct DEP exposure were not observed. Other *high* or *medium confidence* studies that evaluated gross and histopathological alterations in ovaries, uteri, vaginas, or mammary glands did not observe DEP-induced effects (Fujii et al. 2005; Procter & Gamble 1994; Shiraishi et al. 2006). Since effects were only observed in one *low confidence* study, the evidence for histopathological effects on female reproductive organs is considered *indeterminate*.

In adolescent F1 female rats exposed to DEP at 1375 mg/kg-day, the onset of puberty as measured by the age at vaginal opening was delayed by 6% compared to control (Fujii et al. 2005). Animals in this dose group had decreased growth in early postnatal life but had reached similar body weights compared to controls at the time puberty was attained, which suggests that the delay in puberty was related to delayed growth. AGD in F1 or F2 female pups was not affected in this study (Fujii et al. 2005) or in F1 females in the study by Gray et al. (2000). There were no effects on estrous cyclicity in F0 or F1 females (Fujii et al. 2005). Likewise, Shiraishi et al. (2006) evaluated serum hormone levels in adult female rats exposed to DEP for 28 days and found no effects on estrous cycles measured during the last week of the 28-day exposure period, but this outcome was considered *low confidence* due to a short observation duration and lack of quantitative data. Taken together, evidence for effects on female morphological development, estrous cyclicity, and hormones were all considered to be *indeterminate*.

3.5.2. Synthesis of results for female reproductive effects—The available studies suggest there is *slight* evidence that DEP is a female reproductive toxicant (Table 3). Some statistically significant effects were reported for decreased gestation length, decreased litter size, maternal body weight gain, organ weight changes, and age at puberty from two *high confidence* multi-generational studies in rats and mice. Effects on gestation length (Fujii et al. 2005) and litter size (RTI International 1984) in these studies were observed in F1 parental animals but not in F0, possibly suggesting that the F1 animals may have increased sensitivity due to their developmental exposure to DEP; however, in all cases, the magnitude of effect was small and was observed only at high dose levels. Otherwise, except for some organ weight and histopathological changes in *low confidence* studies, results across studies were largely negative. It is possible that differences in test animal

species/strains and experimental designs may contribute to some of these inconsistencies across studies.

3.6. Developmental effects

Figs. indicating the doses at which statistically significant developmental effects occurred are provided in the Supplementary Materials (Figs. S8–S12 and S18). The effects described in this section are limited to survival, growth, and structural alterations; other effects in developing animals (e.g. reproductive effects) are described in respective sections.

3.6.1. Summary of available studies—Four studies evaluated fetal survival following gestational exposure to DEP (Furr et al. 2014; Howdeshell et al. 2008; NTP 1988; Procter & Gamble 1994), and five studies evaluated the number of live pups at birth following gestational (Gray et al. 2000; Hardin et al. 1987; Setti Ahmed et al. 2018) or multigenerational (Fujii et al. 2005; RTI International 1984) exposure to DEP. Of these, only the continuous breeding study in CD-1 mice (RTI International 1984) observed a dose-related effect. The number of live F2 offspring at birth (males and females combined) was statistically significantly decreased by 14% in the 3640 mg/kg-day DEP exposure group relative to controls, whereas no effects on survival were observed in the F1 offspring. In a gestational exposure study in Sprague-Dawley rats, Howdeshell et al. (2008) reported a statistically significant increase in resorptions and fetal mortality at 600 mg/kg-day DEP, but this effect was not observed in any other higher or lower DEP dose groups and thus did not appear to be treatment-related. Otherwise, in the rat two-generation study and in the remaining gestational exposure studies in rats, mice, and rabbits, there were no effects on the number of implantations or resorptions (Fujii et al. 2005; NTP 1988; Procter & Gamble 1994), viability of fetuses (Furr et al. 2014; NTP 1988; Procter & Gamble 1994), or viability of pups at birth (Fujii et al. 2005; Gray et al. 2000; Hardin et al. 1987; Setti Ahmed et al. 2018). There was no effect on offspring sex ratio in any of the studies that assessed this endpoint, which included studies in rats, mice, and rabbits (Fujii et al. 2005; NTP 1988; Procter & Gamble 1994; RTI International 1984). Given the evidence of decreased fetal survival in the high confidence continuous breeding study in mice but lack of effect in all other studies, the evidence for DEP effects on fetal survival following gestational exposure is considered indeterminate.

The three studies that reported postnatal survival found differing results. In F1 female rats exposed to a low dose of DEP (0.173 mg/kg-day) from birth (PND 1) through sacrifice, there was a statistically significant increase in the postnatal mortality of F2 pups through PND 14 (Manservisi et al. 2015), although this result is considered *low confidence* due to the concerns discussed in Section 3.3. In contrast, no significant effects were observed in *high confidence* studies. The two-generation study in rats from Fujii et al. (2005) indicated a non-significant trend towards lower survival of F1 offspring at PND 21 in the DEP treatment groups, but this trend was not observed in F2 offspring and it is not clear that it is treatment-related. There was no effect on F1 offspring viability through PND 3 in mice exposed to 4,500 mg/kg-day DEP from GD 6–13 (Hardin et al. 1987). Taken together, the evidence for the effects of DEP on postnatal survival is considered *indeterminate*.

No effects on fetal growth were observed in seven studies conducted in rats, mice, and rabbits, of which all but one were considered *high confidence* for this outcome. The two-generation reproductive study in rats (Fujii et al. 2005) and the continuous breeding study in mice (RTI International 1984) both reported that F1 and F2 offspring body weights at birth were similar between DEP-treated animals and controls. Likewise, four gestational exposure studies indicated that fetal body weights of rats and rabbits (NTP 1988; Procter & Gamble 1994) or body weights of rats or mice at birth (Gray et al. 2000; Hardin et al. 1987; Setti Ahmed et al. 2018) were similar between DEP-treated animals and controls. While this data suggests that DEP does not impact fetal growth across multiple species, it was decided among reviewers that there were not enough studies available to support a judgement of *compelling evidence of no effect.* The evidence for effects of DEP on prenatal growth is therefore considered *indeterminate*.

Conversely, decreased postnatal growth was reported in multiple studies that evaluated this endpoint. In the *high confidence* two-generation reproduction study, Fujii et al. (2005) observed that F1 and F2 Sprague-Dawley rat offspring exposed to 1,297 mg/kg-day and 1,375 mg/kg-day DEP, respectively, had lower body weights relative to controls at PND 4, 7, 14, and 21. The decrease in body weight was statistically significant for F1 female pups at all time points, and for F1 and F2 male pups and F2 female pups at PND 21. These pups also had a delay in pinna detachment (a developmental biomarker) relative to controls, which was statistically significant for F1 males. Similarly, the high confidence continuous breeding study in mice by RTI International (1984) found that F1 male and female pup body weights at the time of weaning (PND 21) were lower in DEP-treated groups (3,640 mg/kg-day) compared to controls and remained lower than controls at the time of mating (PND 74 \pm 10), although the authors did not perform a statistical analysis. Additionally, the low confidence multigenerational studies in rats observed statistically significant reductions in F1 weanling body weights (Pereira and Rao 2007), F1 adult body weights (Hu et al. 2016), and F2 body weights at PNDs 7 and 14 (Manservisi et al. 2015). The low confidence study by Setti Ahmed et al. (2018) that exposed rats from GD 8 – PND 30 reported a statistically significant decrease in F1 body weights beginning at PND 15. There was no effect on F1 offspring growth in the gestational exposure studies in mice by Hardin et al. (1987) (evaluated through PND 3) and in rats by Gray et al. (2000) (evaluated as adults), although these studies used shorter exposure duration (GD 6-13 and GD 14-PND 3, respectively). In peripubertal males exposed to DEP for 7 days, no effects on body weight were observed (Oishi and Hiraga 1980). Given the consistent reductions in offspring postnatal growth across all multigenerational studies including the *high confidence* studies by Fujii et al. (2005) and RTI International (1984), there is robust evidence that DEP exposure can affect postnatal growth.

The two *high confidence* studies that examined the incidence of fetal external, skeletal, and visceral malformations in rats (NTP 1988) and rabbits (Procter & Gamble 1994) found little to no effects of DEP exposure during gestation. However, the study in rats by NTP (1988) observed a dose-related statistically significant increase in the mean percent of fetuses with a rudimentary or full extra (supernumerary) lumbar rib, which is a skeletal variation (US EPA 1991). The study by Procter & Gamble (1994) in rabbits observed two malformed fetuses in two different litters in the highest DEP dose group (out of 80 total fetuses and

12 total litters in this dose group) and no malformed fetuses in the controls or lower DEP dose groups (out of 77–91 total fetuses and 12 litters per group); the observed malformations consisted of one fetus with fused and split ribs and missing lumbar and coccygeal vertebrae, and one with acrania, hernia umbilicalis, and incurved ribs. The authors did not consider this finding to be treatment-related because the malformations were of different types and the incidence was within the rate of historical controls, although historical control data were not provided as part of this study. There were no dose-related effects on skeletal variations in the rabbit fetuses, including on the number of ribs. Additionally, the *low confidence* study by Setti Ahmed et al. (2018) evaluated the morphological development of the intestines in F1 rats exposed from GD 8 – PND 30 and found alterations in enterocytes as well as supporting mechanistic evidence indicating decreased cell proliferation and enzyme activities but did not provide any quantitative data. Altogether, the evidence for DEP effects on fetal morphological development is considered *slight*.

3.6.2. Synthesis of results for developmental effects—Overall, there is *moderate* evidence that DEP exposure can cause developmental toxicity (Table 4). In particular, decreased postnatal growth in rats and mice was observed in almost all gestational and early postnatal exposure studies that assessed this endpoint, including two studies that were *high confidence* for this endpoint (Fujii et al. 2005; RTI International 1984). Otherwise, effects on survival, growth, and structural alterations were generally not observed. It is notable that in many cases, the observed developmental effects occurred at doses that were also associated with maternal effects (Section 3.5). For instance, fetal skeletal variations occurred at a dose concurrent with decreased maternal body weight gain (NTP 1988), and decreased offspring postnatal growth occurred at doses associated with decreased litter size (RTI International 1984). This indicates that DEP is not a potent developmental toxicant, even though multiple developmental effects were observed.

Although the number of studies was limited, the available body of literature for developmental DEP exposure includes a variety of experimental designs covering exposure during critical windows of fetal and postnatal development, including two studies that examined outcomes in offspring following exposure over two generations (Fujii et al. 2005; RTI International 1984). Another strength is that several species were examined in these studies.

3.7. Liver effects

A figure indicating the doses at which statistically significant effects on liver occurred is provided in the Supplementary Materials (Fig. S13–S15 and S18).

3.7.1. Summary of available studies—Increases in liver weight in rats and mice were reported in multiple *high* and *medium confidence* studies that tested oral and dermal exposures of DEP. Statistically significant increases in relative liver weight ranged from 7 to 33%, and were generally observed at the highest dose tested (> 1000 mg/kg-day) (Brown et al. 1978; Fujii et al. 2005; Moody and Reddy 1978; Oishi and Hiraga 1980; RTI International 1984). In a 2-year dermal exposure study, relative liver weights were statistically significantly increased following 4 weeks of exposure in rats and mice, but

not at the 15-month evaluation (NTP 1995). One potential mechanism for the increase in liver weight may be the induction of peroxisome proliferation. Fujii et al. (2005) reported a statistically significant increase at > 1150 mg/kg-day in the microsomal Cyp4A1 isoenzyme induced in rodents by peroxisome proliferators. Peroxisome proliferation was only slightly increased in rat hepatocytes following exposure to DEP despite a statistically significant increase in peroxisomal associated enzyme carnitine acetyltransferase and relative liver weight at 1753 mg/kg-day (Moody and Reddy 1978). In contrast, in several high confidence studies in rats, there was no effect on absolute liver weight following gestational exposure in F1 males at 750 mg/kg-day DEP (Gray et al. 2000) and no effect on relative liver weight in adult animals following 2 or 4-week exposures to 500 mg/kg-day DEP or 250 mg/kg-day MEP (Kwack et al. 2009; Kwack et al. 2010) or up to 1000 mg/kg-day DEP (Shiraishi et al. 2006). The lack of effect on organ weight may be due to differences in the exposure window and/or dose. A low confidence study reported that absolute liver weight was decreased in rat pups at PND 15 and 30 (Setti Ahmed et al. 2018). In a series of low dose (0.57–6.25 mg/ kg-day) studies, effects on liver weights were inconsistent, with studies reporting increases, decreases, or no effect on liver weight following oral DEP exposures from 90 to 150 days (Mapuskar et al., 2007; Pereira et al., 2006, 2007a, 2007c, 2008b, Pereira and Rao, 2006a, 2007; Sonde et al., 2000). These low dose studies were considered low confidence due to the concerns discussed in Section 3.3. Despite these inconsistencies, increased liver weights were observed across multiple high and medium confidence studies, and the evidence for changes in liver weight following DEP exposure is considered moderate.

Data on histopathological changes differed across studies. As noted above, the *medium* confidence study in rats by Moody and Reddy (1978) reported a slight increase in hepatic peroxisome proliferation, measured as the mitochondrial-peroxisome ratio using electron microscopy. No dose-related histopathological changes in the liver were observed in rats of either sex following 4 weeks of exposure to doses up to 1000 mg/kg-day (Shiraishi et al. 2006), or in F0 and F1 rats in a multigenerational reproductive study at doses up to 1016-1297 mg/kg-day (F0) and 1150-1375 mg/kg-day (F1) (Fujii et al. 2005), both of which are considered *medium confidence* studies. Similarly, except for a statistically significant increased incidence of basophilic foci in the liver of male mice at 16.8 mg/kgday, no dose-related histopathological changes in liver were reported in the high confidence two-year dermal exposure study in mice at doses up to 33.6 mg/kg-day (NTP 1995). No histopathological changes in liver were observed in either sex following 42 or 112 days of exposure of rats to doses up to 3160 mg/kg-day in the low confidence study by Brown et al. (1978). Conversely, histopathological changes including vacuolization, lipid droplets in the liver, loss of hepatic architecture and necrotic changes were reported in the series of low dose (0.57-6.25 mg/kg-day) studies considered low confidence due to the concerns discussed in Section 3.3 (Mapuskar et al., 2007; Pereira et al., 2006, 2007a, 2007c, 2008b, Pereira and Rao, 2006a, 2007; Sinkar and Rao, 2007). The histopathological reports from this group did not include quantitative data, which contributed to the low confidence in these findings. Given the lack of confidence in the low dose findings along with the negative findings at higher doses over various periods of exposure, the evidence for histopathological effects is considered indeterminate.

Four high confidence studies reported limited evidence of effects on biochemical markers of hepatocellular or hepatobiliary liver toxicity following DEP exposure. Statistically significantly decreased aspartate aminotransferase (AST) activity, a non-specific marker of liver injury, was reported in male rats following exposure to up to 1000 mg/kg-day for 28 days by Shiraishi et al. (2006). The biological significance of decreased AST activity is unclear. Statistically significant increases in a hepatobiliary marker, gamma-glutamyl transferase (GGT), were also reported in this study at 40 and 1000 mg/kg-day. Findings in rats exposed to 500 mg/kg-day DEP or 250 mg/kg-day MEP for 2 or 4 weeks showed only minimal (0–21%) non-statistically significant increases in alanine aminotransferase (ALT), a liver specific marker, or AST activity in serum compared to controls (Kwack et al. 2009; Kwack et al. 2010). Additionally, a statistically significant increase in alkaline phosphatase (ALP) activity, also marker of hepatobiliary injury, was observed in female rats following a 15-month dermal exposure, with no effects observed in male rats (NTP 1995). In contrast, statistically significant changes in the serum levels of ALT, AST, ALP, sorbitol dehydrogenase (SDH), and lactate dehydrogenase (LDH) were frequently reported in rats and mice in the series of low dose studies considered *low confidence* due to the concerns discussed in Section 3.3 (Mapuskar et al., 2007; Pereira et al., 2006, 2007a, 2007c, 2008, Pereira and Rao, 2006a, 2006b, 2007; Sinkar and Rao, 2007; Sonde et al., 2000). Changes in enzyme levels were observed in both sexes and across generations; however, the number of animals utilized to assess enzymes changes was typically small (n

6) and the findings were not always consistent. For example, one low dose study from this group reported decreased serum LDH activity levels in male and female rats without changes in serum levels of AST, SDH or ALP (Sinkar and Rao 2007). Other statistically significant but inconsistent changes in clinical chemistry in the low confidence low dose studies were altered serum triglycerides, serum glucose and liver glycogen levels in mice and rats (Mapuskar et al., 2007; Pereira et al., 2006, 2007a, 2007b, 2008, Pereira and Rao, 2006a, 2006b; Sonde et al., 2000); altered cholesterol in rats (Mapuskar et al., 2007; Pereira et al., 2006, 2007a, 2007b, 2008, Pereira and Rao, 2006a, 2006b; Sonde et al., 2000); and increased lipid peroxidation and decreased glutathione and glutathione reductase levels (Pereira et al., 2006, 2007a, 2007b, 2008, Pereira and Rao, 2006a, 2006b; Sonde et al., 2000). In one *high confidence* study, a statistically significant increase in the fatty acid oxidizing enzyme CYP4A was reported at 1150 mg/kg-day, supportive of alterations in lipid metabolism (Fujii et al. 2005). In addition, in one *medium confidence* study the activity of two peroxisome associated enzymes, catalase and carnitine acetyltransferase, was statistically significantly increased in male rats following subchronic exposure (Moody and Reddy 1978), and another *medium confidence* study of the same design reported statistically significantly decreased serum triglycerides and no change in serum cholesterol (Moody and Reddy 1982). Overall, statistically significant increases in liver enzyme activity indicative of liver toxicity were observed mostly in low confidence studies; therefore, the evidence for effects on biochemical markers of liver toxicity is slight.

3.7.2. Synthesis of results for liver effects—The available toxicology studies in rodents provide *moderate* evidence of liver toxicity following DEP exposure (Table 5). Evidence for liver toxicity in experimental animal studies includes reports on liver weight and clinical chemistry, as well as histopathology findings. Statistically significant increases

in liver weight were demonstrated in several *high, medium*, and *low confidence* studies of rats and mice across multiple routes of exposure, with effects observed more consistently at higher doses in *high* and *medium confidence* studies. Histopathological and biochemical changes were frequently reported in a series of *low confidence*, low-dose studies of DEP in rats and mice whereas there was little evidence of significant histopathological or biochemical changes in *high* and *medium* confidence studies. Effects of DEP on liver weight may be due to its actions as a peroxisome proliferator, which is a common mechanism among phthalates; however, DEP is considered a relatively weak peroxisome proliferator (Moody and Reddy 1978; Okita and Okita 1992). Therefore, other adaptive mechanisms may play a role in changes in liver weight at higher exposures of DEP (Boone et al. 2005; Hall et al. 2012; Williams and Iatropoulos 2002).

Strengths of the evidence base for hepatic effects include the availability of a variety of experimental designs covering a range of doses. The most remarkable findings were observed in low dose studies that were considered *low confidence*. Therefore, the database for liver effects would be strengthened by additional research at exposures in the lower dose range using more robust experimental designs.

3.8. Kidney effects

A figure indicating the doses at which statistically significant effects on kidney weight occurred is provided in the Supplementary Materials (Fig. S16).

3.8.1. Summary of available studies—Changes in kidney weight following oral DEP exposure in rats and mice were inconsistent. Multiple high and medium confidence oral exposure studies in rats reported statistically significant increases or decreases in relative and/or absolute kidney weight, with no clear or consistent pattern of effect across studies (Brown et al. 1978; Fujii et al. 2005; Shiraishi et al. 2006). One high confidence dermal study in rats and mice reported a statistically significant increase in relative kidney weight in female mice after 15 months of exposure; however, in male mice, absolute kidney weight was statistically significantly decreased over the same exposure period, and there were no effects on absolute or relative kidney weight in male or female rats in this study (NTP 1995). Dose-related increases or decreases in kidney weight were not observed in rats in three high or medium confidence peripubertal exposure studies (Kwack et al. 2009; Kwack et al. 2010; Oishi and Hiraga 1980) or in the *high confidence* study that exposed rats from GD 14 – PND 3 (Gray et al. 2000). In one low confidence study, in which rat dams were exposed to DEP from GD 8 to PND 30, kidney weight was decreased in pups on PND 15 and PND 30 (Setti Ahmed et al. 2018). Taken together, because of inconsistent changes in kidney weight in mice and rats, the evidence for effects on kidney weight is considered indeterminate.

Histopathological analyses of the kidney were conducted in four of the same studies that evaluated kidney weight in rats or mice (Brown et al. 1978; Fujii et al. 2005; NTP 1995; Shiraishi et al. 2006). In all cases, no dose-related histopathological changes of the kidney were reported. Altogether, given the lack of effect across studies, the evidence for histopathological effects in the kidney is considered *indeterminate*.

A statistically significant increase in serum calcium levels compared to controls was observed in rats exposed to 500 mg/kg-day for 2 or 4 weeks in the studies by Kwack et al. (2010) and Kwack et al. (2009), respectively; whereas, other biochemical parameters related to the kidney were not affected in these animals (urinary protein level and serum creatinine, creatinine kinase, blood urea nitrogen, and albumin). In the study in rats by Brown et al. (1978), effects on urinary cell excretion were inconsistent, with a statistically significant decrease observed in males after 13 weeks of exposure to 3,160 mg/kg-day but otherwise there were no dose or time-related trends. Since the only consistent effect was increased serum calcium levels, the evidence for effects on biochemical markers of kidney damage is considered *indeterminate*.

3.8.2. Synthesis of results for kidney effects—It was concluded that there was *indeterminate* evidence for kidney toxicity following exposure to DEP (Table 6). The most convincing evidence for kidney toxicity comes from *medium* and *high confidence* studies that observed statistically significant increases or decreases in absolute and/or relative kidney weight at higher doses, although these effects on kidney weight were inconsistent both within and across studies. Minimal data were available on biochemical measurements that are indicative of kidney toxicity. Changes in kidney weight were not supported by histopathological changes in male or female rat kidney.

3.9. Cancer

A 2-year dermal bioassay in rats and mice demonstrated an increased incidence of liver neoplasms in mice following dermal DEP exposure (Fig. S17) (NTP 1995). Significant increases in tumor formation in other tissues was not observed in this study. The combined incidence of adenoma or carcinoma was increased in males exposed to 33 mg/day, along with a positive dose-related trend; however, the increase is within NTP's historical control values for 2-year mouse studies. In treated females, combined incidence of adenoma or carcinoma was increased as compared to control but was not related to dose. No evidence of carcinogenic activity of DEP was observed in F344N rats, but sensitivity was reduced due to low survival rates in the rats. In addition, 1-year initiation/promotion studies in CD-1 mice that tested DEP with and without the skin tumor promoter 12-O-tetradecanoylphorbol-13-acetate (TPA) with or without the skin tumor initiator 7,12-dimethylbenz[a]anthrancene (DMBA) demonstrated that DEP was not able to imitate or promote skin neoplasms (NTP 1995). In the absence of increased tumor incidence above background levels and negative findings in the initiation and promotion studies, the evidence of liver tumorgenicity in experimental animals dermally exposed is considered *indeterminate* (Table 7).

4. Discussion

The results of this systematic review provide *moderate* evidence that DEP causes male reproductive toxicity, developmental toxicity, and liver toxicity, and *slight* evidence for female reproductive toxicity. Evidence for effects on kidney toxicity and cancer was considered *indeterminate*. These conclusions are for hazard identification and do not consider the relative sensitivity of each of these outcomes to DEP exposure. By way of comparison, the CHAP review of DEP toxicity compared the lowest-observed-adverse-

effect levels among a smaller subset of DEP studies and found that liver weight was the most sensitive systemic endpoint (Consumer Product Safety Commission 2011). The CHAP also concluded that DEP was not a potent developmental toxicant, although they found that developmental endpoints were more sensitive than reproductive endpoints in the multigenerational studies by Fujii et al. (2005) and RTI International (1984).

There were several outcomes for which no effects were observed, most notably fetal testosterone production and fetal growth. However, in no case was there found to be enough evidence to warrant a conclusion of *compelling evidence of no effect*. A judgement of *compelling evidence of no effect* represents an uncommon situation where no association was identified from extensive evidence across a range of populations and exposures. In the case of fetal testosterone and other androgen-dependent developmental outcomes, all available studies were conducted in the same species (rat) and used mostly similar experimental designs, whereas studies in multiple species and exposure scenarios would be necessary to rule out the possibility of effect. For fetal growth, studies were available from three species (rats, mice, rabbits) and multiple exposure scenarios (gestation-only and multigenerational studies), but the number of *high confidence* studies (six) was still relatively low. Overall, a larger database of studies for DEP demonstrating a lack of effect would be necessary in order to have confidence in a conclusion of *compelling evidence of no effect* for any of these endpoints.

Although DEP induced some developmental outcomes in rats and mice (decreased postnatal growth, increased supernumerary ribs), there were no effects on fetal growth, and only one study indicated effects on fetal survival (RTI International 1984). Other developmentally toxic phthalates have been found to decrease fetal survival, and it has been hypothesized that phthalate-induced decreases in fetal testosterone production and fetal survival may both be caused by decreased steroidogenesis (decreased testicular testosterone production in male fetuses and decreased ovarian progesterone production in dams) (Howdeshell et al. 2008). It is therefore plausible that the lack of effect of DEP on fetal survival is related to its low potency as a disruptor of steroidogenesis, although further research would be necessary to evaluate this hypothesis.

There is some support in the epidemiological literature for an association between DEP exposure and reproductive and developmental outcomes. A systematic review of epidemiological literature by our colleagues found *slight* evidence of an association between DEP exposure and male reproductive outcomes in humans. Inconsistent associations were found between DEP exposure and testosterone or AGD despite the relatively high exposure levels, and it was concluded that DEP does not appear to have a strong antiandrogenic effect in humans. There was also some support in the epidemiological literature for an inverse association between DEP exposure and semen quality, although most studies observed no association or a positive association (Radke et al. 2018). Interestingly, systematic review of the epidemiological literature on the female reproductive effects of phthalates found multiple studies that reported an association with early onset of puberty, pregnancy loss, and preterm birth (Radke et al. 2019b), which are similar to some of the outcomes seen in the animal studies.

This systematic review was limited to six hazard categories, which we selected a priori due to the likelihood that they would have sufficient data available to inform hazard evaluation. We found that there were multiple studies available for most of the hazard categories, although only one study investigated cancer. We also observed that few studies evaluated effects for DEP that were not captured by the PECO. In particular, effects related to neurological development were evaluated in several of the included studies: Fujii et al. (2005) performed daily reflex response tests on F1 and F2 Sprague-Dawley rat pups and observed no significant differences between groups, whereas Pereira and Rao (2007) noted that pups exposed to DEP had sluggish behavior and movement reduction compared to controls but did not provide quantitative data, and Setti Ahmed et al. (2018) reported a slight decrease in F1 brain weight after gestational and lactational exposure to DEP. Additionally, several of the included studies reported pituitary, adrenal, or thyroid weights or histopathology (Fujii et al. 2005; Kwack et al. 2009; NTP 1995; RTI International 1984; Shiraishi et al. 2006), with some dose-related organ weight changes observed, and Shiraishi et al. (2006) reported no effects on thyroid hormone levels in male or female rats after 28 days of exposure. Two excluded studies by Pereira et al. (2007b, 2008c) reported degenerative effects on the adrenal cortex and thyroid after multigenerational exposure to low doses of DEP. Although we did not formally evaluate the studies by Pereira et al. (2007b, 2008c) because they did not report any of the outcomes listed in the PECO, they appear to have been conducted on the same animals used in other studies by this group and have the same concerns for risk of bias and sensitivity, so it is likely that these findings would be considered low confidence if they had been evaluated. Systematic review of emerging outcomes in the epidemiological literature found *slight* evidence of an association between DEP exposure and metabolic effects (insulin resistance and blood glucose/impaired glucose tolerance) (Radke et al., 2019a), and limited support for an association between DEP exposure and neurodevelopmental effects (Radke et al., 2020).

This systematic review highlighted several ways in which future studies could provide further insight into the mechanisms and characterization of hazards of DEP exposure. Most of the available studies were conducted at doses that are far greater than doses expected to be relevant to human exposures, which are generally estimated to be in the µg/kg-day range (Consumer Product Safety Commission 2014). The available low dose studies come closer to recapitulating human exposure levels but had significant concerns raised during study evaluation and reported findings that were frequently not supported by studies that tested higher dose levels. The recent review of DEP toxicity by the Consumer Product Safety Commission (2011) reached a similar conclusion that the studies by Pereira and coauthors had questionable reliability and were not consistent with effects observed in other studies. Additional studies are needed to confirm findings by these groups that have shown dramatic effects following chronic exposure to DEP at low doses.

Additionally, whereas there is little evidence of DEP-mediated effects on testosterone, an interesting finding of this systematic review is that effects on sperm were observed in several studies that exposed peripubertal or adult animals to DEP for longer durations. This finding is suggestive of the effects on Sertoli cells, seminiferous tubules, and germ cell development that have been observed for other phthalates and are thought to be mediated through an androgen-independent mode of action. For DEP, however, the magnitude of effect on sperm

parameters was relatively low and varied across the few available studies. Further studies may be warranted to determine the extent to which this mode of action is conserved for DEP, which has the potential to lead to male reproductive effects in the absence of effects on steroidogenesis.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgement:

This manuscript is dedicated to the memory of Raghu Nath, a biologist at the EPA National Center for Environmental Assessment, who contributed to this systematic review of DEP. We would like to acknowledge Anna Chen, Evangela Matthews, Swati Gummadi, Carolyn Gigot, Mefruz Haque, and Andrew Greenhalgh (EPA student services contractors) for their assistance in data extraction and visualization for this systematic review.

References

- Bailey SA, Zidell RH, Perry RW, 2004. Relationships between organ weight and body/brain weight in the rat: What is the best analytical endpoint? Toxicol. Pathol 32, 448–466. [PubMed: 15204968]
- Boone L, Meyer D, Cusick P, Ennulat D, Bolliger AP, Everds N, et al., 2005. Selection and interpretation of clinical pathology indicators of hepatic injury in preclinical studies. Vet. Clin. Pathol 34, 182–188. [PubMed: 16134065]
- Brown D, Butterworth KR, Gaunt IF, Grasso P, Gangolli SD, 1978. Short-term oral toxicity study of diethyl phthalate in the rat. Food Cosmet. Toxicol 16, 415–422. [PubMed: 711065]
- Clark KE, David RM, Guinn R, Kramarz KW, Lampi MA, Staples CA, 2011. Modeling human exposure to phthalate esters: A comparison of indirect and biomonitoring estimation methods. Hum. Ecol. Risk Assess 17, 923–965. [PubMed: 23087593]
- Clewell RA, Campbell JL, Ross SM, Gaido KW, Clewell HJ III, Andersen ME, 2010. Assessing the relevance of in vitro measures of phthalate inhibition of steroidogenesis for in vivo response. Toxicol. In Vitro 24, 327–334. [PubMed: 19699292]
- Consumer Product Safety Commission, 2011. Final toxicity review for diethyl phthalate (DEP)
- Consumer Product Safety Commission, 2014. Chronic Hazard Advisory Panel on phthalates and phthalate alternatives (with appendices) MD:U.S. Consumer Product Safety Commission, Directorate for Health Sciences, Bethesda.
- Cooper G, Lunn R, Agerstrand M, Glenn B, Kraft A, Luke A, et al. , 2016. Study sensitivity: Evaluating the ability to detect effects in systematic reviews of chemical exposures. Environ. Int 92–93, 605–610.
- Field EA, Price CJ, Sleet RB, George JD, Marr MC, Myers CB, et al., 1993. Developmental toxicity evaluation of diethyl and dimethyl phthalate in rats. Teratology 48, 33–44. [PubMed: 8351646]
- Foster PMD, Thomas LV, Cook MW, Walters DG, 1983. Effect of di-n-pentyl phthalate treatment on testicular steroidogenic enzymes and cytochrome P-450 in the rat. Toxicol. Lett 15, 265–271. [PubMed: 6829050]
- Fujii S, Yabe K, Furukawa M, Hirata M, Kiguchi M, Ikka T, 2005. A two-generation reproductive toxicity study of diethyl phthalate (DEP) in rats. J. Toxicol. Sci 30, S97–116.
- Furr JR, Lambright CS, Wilson VS, Foster PM, Gray LE Jr., 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. [PubMed: 24798384]
- Gray LE Jr., Ostby J, Furr J, Price M, Veeramachaneni DNR, 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. [PubMed: 11099647]

- Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al., 2011. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. J. Clin. Epidemiol 64, 383–394. [PubMed: 21195583]
- Hall AP, Elcombe CR, Foster JR, Harada T, Kaufmann W, Knippel A, et al., 2012. Liver hypertrophy: A review of adaptive (adverse and non-adverse) changes–conclusions from the 3rd International ESTP Expert Workshop. Toxicol. Pathol 40, 971–994. [PubMed: 22723046]
- Hardin BD, Schuler RL, Burg JR, Booth GM, Hazelden KP, Mackenzie KM, et al., 1987. Evaluation of 60 chemicals in a preliminary developmental toxicity test. Teratog. Carcinog. Mutagen 7, 29– 48. [PubMed: 2884741]
- Haseman JK, Bailer AJ, Kodell RL, Morris R, Portier K, 2001. Statistical issues in the analysis of low-dose endocrine disruptor data. Toxicol. Sci 61, 201–210. [PubMed: 11353128]
- Hazleton Laboratories America Inc, 1983a. Screening of priority chemicals for potential reproductive hazard (final report) with attachments and cover sheet EPA/OTS Doc #86–870001624. Shell Oil Company.
- Hazleton Laboratories America Inc, 1983b. Final report. Screening of priority chemicals for potential reproductive hazard NIOSH/00133459. Cincinnati, OH: NIOSH (The National Institute for Occupational Safety and Health).
- Hazleton Laboratories America Inc, 1992. Initial submission: Screening of priority chemicals for potential reproductive hazard (final report) with cover letter dated 041492. 88920002015. Washington, DC: U.S. Environmental Protection Agency.
- Higgins J, 2011. Cochrane handbook for systtematic reviews of interventions In: Chapter 8: Assessing risk of bias in included studies The Cochrane collaboration [updated March 2011].
- Hill AB, 1965. The environment and disease: Association or causation? Proc. R. Soc. Med 58, 295–300. [PubMed: 14283879]
- Howdeshell KL, Wilson VS, Furr J, Lambright CR, Rider CV, Blystone CR, et al., 2008. A mixture of five phthalate esters inhibits fetal testicular testosterone production in the Sprague-Dawley rat in a cumulative, dose-additive manner. Toxicol. Sci 105, 153–165. [PubMed: 18411233]
- Hu G, Li J, Shan Y, Li X, Zhu Q, Li H, et al., 2018. In utero combined di-(2-ethylhexyl) phthalate and diethyl phthalate exposure cumulatively impairs rat fetal Leydig cell development. Toxicology 395, 23–33. [PubMed: 29325824]
- Hu J, Raikhel V, Gopalakrishnan K, Fernandez-Hernandez H, Lambertini L, Manservisi F, et al. 2016. Effect of postnatal low-dose exposure to environmental chemicals on the gut microbiome in a rodent model 4, 26.
- Johnson KJ, Heger NE, Boekelheide K, 2012. Of mice and men (and rats): Phthalate-induced fetal testis endocrine disruption is species-dependent. Toxicol. Sci 129, 235–248. [PubMed: 22700540]
- Jones HB, Garside DA, Liu R, Roberts JC, 1993. The influence of phthalate esters on Leydig cell structure and function in vitro and in vivo. Exp. Mol. Pathol 58, 179–193. [PubMed: 8519345]
- Kwack S, Kim K, Kim H, Lee B, 2009. Comparative toxicological evaluation of phthalate diesters and metabolites in Sprague-Dawley male rats for risk assessment. J. Toxicol. Environ. Health Part A 72, 1446–1454.
- Kwack SJ, Han EY, Park JS, Bae JY, Ahn IY, Lim SK, et al., 2010. Comparison of the short term toxicity of phthalate diesters and monoesters in Sprague-Dawley male rats. Toxicol. Res 26, 75– 82. [PubMed: 24278509]
- Lamb J, Chapin R, Teague J, Lawton A, Reel J, 1987. Reproductive effects of four phthalic acid esters in the mouse. Toxicol. Appl. Pharmacol 88, 255–269. [PubMed: 3564043]
- Liu K, Lehmann KP, Sar M, Young SS, Gaido KW, 2005. Gene expression profiling following in utero exposure to phthalate esters reveals new gene targets in the etiology of testicular dysgenesis. Biol. Reprod 73, 180–192. [PubMed: 15728792]
- Makris SL, Thompson CM, Euling SY, Selevan SG, Sonawane B, 2008. A lifestage-specific approach to hazard and dose-response characterization for children's health risk assessment. Birth Defects Res. B Dev. Reprod. Toxicol 83, 530–546. [PubMed: 19085945]
- Mangelsdorf I, Buschmann J, Orthen B, 2003. Some aspects relating to the evaluation of the effects of chemicals on male fertility. Regul. Toxicol. Pharm 37, 356–369.

- Manservisi F, Gopalakrishnan K, Tibaldi E, Hysi A, Iezzi M, Lambertini L, et al., 2015. Effect of maternal exposure to endocrine disrupting chemicals on reproduction and mammary gland development in female Sprague-Dawley rats. Reprod. Toxicol 54, 110–119. [PubMed: 25554385]
- Mapuskar K, Pereira C, Rao CV, 2007. Dose-dependent sub-chronic toxicity of diethyl phthalate in female Swiss mice. Pestic. Biochem. Physiol 87, 156–163.
- Martino-Andrade AJ, Chahoud I, 2010. Reproductive toxicity of phthalate esters. Mol. Nutr. Food Res 54, 148–157. [PubMed: 19760678]
- Mondal S, Ghosh S, Bhattacharya S, Mukherjee S, 2019. Chronic dietary administration of lower levels of diethyl phthalate induces murine testicular germ cell inflammation and sperm pathologies: Involvement of oxidative stress. Chemosphere 229, 443–451. [PubMed: 31085347]
- Moody D, Reddy J, 1978. Hepatic peroxisome (microbody) proliferation in rats fed plasticizers and related compounds. Toxicol. Appl. Pharmacol 45, 497–504. [PubMed: 705785]
- Moody DE, Reddy JK, 1982. Serum triglyceride and cholesterol contents in male rats receiving diets containing plasticizers and analogues of the ester 2-ethylhexanol. Toxicol. Lett 10, 379–383. [PubMed: 7089991]
- National Research Council, 2008. In: Phthalates and cumulative risk assessment: The tasks ahead Washington (DC): National Academies Press (US). Copyright 2008 by the National Academy of Sciences. All rights reserved.
- National Academies of Sciences, Engineering, and Medicine, 2018. Progress toward transforming the Interated Risk Information System (IRIS) program. A 2018 evaluation The National Academies Press, Washington, DC.
- NTP, 1988. Developmental toxicity evaluation of diethyl phthalate (CAS no. 84–66-2) administered to CD rats on gestational days 6 through 15. NTP- 88–336, RTI–207
- NTP, 1995. Toxicology and carcinogenesis studies of diethylphthalate (CAS no. 84–66-2) in F344/N rats and B6C3F1 mice (dermal studies) with dermal initiation/promotion study of diethylphthalate and dimethylphthalate (CAS no. 131–11-3) in male Swiss (CD-1(R)) mice. NTP TR 429 Research Triangle Park, NC.
- Oishi S, Hiraga K, 1980. Testicular atrophy induced by phthalic acid esters: Effect on testosterone and zinc concentrations. Toxicol. Appl. Pharmacol 53, 35–41. [PubMed: 7385236]
- Okita RT, Okita JR, 1992. Effects of diethyl phthalate and other plasticizers on laurate hydroxylation in rat liver microsomes. Pharm. Res 9, 1648–1653. [PubMed: 1488412]
- Pereira C, Mapuskar K, Rao CV, 2006. Chronic toxicity of diethyl phthalate in male Wistar rats–a dose-response study. Regul. Toxicol. Pharm 45, 169–177.
- Pereira C, Mapuskar K, Vaman Rao C., 2007a. A two generation chronic mixture toxicity study of Clophen A60 and diethyl phthalate after generational and lactational exposure in female Wistar rats. Pestic. Biochem. Physiol 88, 156–166.
- Pereira C, Mapuskar K, Rao CV, 2008b. Chronic mixture toxicity study of Clophen A60 and diethyl phthalate in male rats. Toxicol. Environ. Chem 90, 349–359.
- Pereira C, Mapuskar K, Vaman Rao C., 2007b. A two-generation chronic mixture toxicity study of Clophen A60 and diethyl phthalate on histology of adrenal cortex and thyroid of rats. Acta Histochem 109, 29–36. [PubMed: 17113135]
- Pereira C, Rao CV, 2006a. Combined and individual administration of diethyl phthalate and polychlorinated biphenyls and its toxicity in female Wistar rats. Environ. Toxicol. Pharmacol 21, 93–102. [PubMed: 21783644]
- Pereira C, Rao CV, 2006b. Chronic toxicity of diethyl phthalate and polychlorinated biphenyls in rats–a sex related biochemical interaction study. Toxicol. Int 13, 53–60.
- Pereira C, Mapuskar K, Rao CV, 2007c. Chronic toxicity of diethyl phthalate–a three generation lactational and gestational exposure study on male Wistar rats. Environ. Toxicol. Pharmacol 23, 319–327. [PubMed: 21783775]
- Pereira C, Rao CV, 2007. Toxicity study of maternal transfer of polychlorinated biphenyls and diethyl phthalate to 21-day-old male and female weanling pups of Wistar rats. Ecotoxicol. Environ. Saf 68, 118–125. [PubMed: 16814384]

- Pereira C, Mapuskar K, Rao CV, 2007d. Reproductive failure associated with chronic interactive mixture toxicity of diethyl phthalate and Clophen A60 after gestational and lactational exposure over two generations in Wistar rats. Toxicol. Int 14, 111–122.
- Pereira C, Mapuskar K, Rao CV, 2008a. Effect of diethyl phthalate on rat testicular antioxidant system: A dose-dependent toxicity study. Pestic. Biochem. Physiol 90, 52–57.
- Pereira C, Mapuskar K, Rao CV, 2008c. A three-generation toxicity study of diethyl phthalate on histology of adrenal and thyroid glands of rats 15, 63–67.
- Procter Gamble, 1994. Teratogenicity study of e-2426.01 (diethyl phthalate) by dermal application to rabbits with cover letter dated 05/02/94. 8EHQ-86940000362 Washington, DC: U.S. Environmental Protection Agency.
- Radke EG, Braun JM, Meeker JD, Cooper GS, 2018. Phthalate exposure and male reproductive outcomes: A systematic review of the human epidemiological evidence. Environ. Int 121, 764– 793. [PubMed: 30336412]
- Radke EG, Galizia A, Thayer KA, Cooper GS, 2019a. Phthalate exposure and metabolic effects: A systematic review of the human epidemiological evidence. Environ. Int 132, 104768. [PubMed: 31196577]
- Radke EG, Glenn BS, Braun JM, Cooper GS, 2019b. Phthalate exposure and female reproductive and developmental outcomes: A systematic review of the human epidemiological evidence. Environ. Int 130, 104580. [PubMed: 31351310]
- Radke EG, Braun JM, Nachman RM, Cooper GS, 2020. Phthalate exposure and neurodevelopment: A systematic review and meta-analysis of human epidemiological evidence. Environ. Int 137, 105408. [PubMed: 32045779]
- RTI International, 1984. Diethyl phthalate: Reproduction and fertility assessment in CD-1 mice when administered in the feed. RTI International, Research Triangle Park, NC.
- Schünemann H, Hill S, Guyatt G, Akl EA, Ahmed F, 2011. The GRADE approach and Bradford Hill's criteria for causation. J. Epidemiol. Commun. Health 65, 392–395.
- Setti Ahmed K, Kharoubi O, Aoues AEK, Bouchekara M, Khaladi B, Taleb M, 2018. Effect of gestational and lactational exposure to DEHP, DINP, and DEP on intestinal morphology, disaccharidases, and alkaline phosphatase in rats during postnatal development. Am. J. Perinatol 35, 1251–1259. [PubMed: 29715699]
- Sharpe RM, 2010. Environmental/lifestyle effects on spermatogenesis. Philos. Trans. R. Soc. Lond. B Biol. Sci 365, 1697–1712. [PubMed: 20403879]
- Shiraishi K, Miyata K, Houshuyama S, Imatanaka N, Umano T, Minobe Y, et al., 2006. Subacute oral toxicity study of diethylphthalate based on the draft protocol for enhanced OECD test guideline no. 407. Arch. Toxicol 80, 10–16. [PubMed: 16059724]
- Singh AR, Lawrence WH, Autian J, 1975. Maternal-fetal transfer of 14c-di-2-ethylhexyl phthalate and 14c-diethyl phthalate in rats. J. Pharm. Sci 64, 1347–1350. [PubMed: 1151708]
- Sinkar PU, Rao CV, 2007. Gender-based comparative toxicity of di-ethyl phthalate in Wistar rats. Toxicol. Environ. Chem 89, 173–183.
- Sonde V, D'souza A, Tarapore R, Pereira L, Khare MP, Sinkar P, et al. , 2000. Simultaneous administration of diethylphthalate and ethyl alcohol and its toxicity in male Sprague-Dawley rats. Toxicology 147, 23–31. [PubMed: 10837929]
- Teitelbaum SL, Li Q, Lambertini L, Belpoggi F, Manservisi F, Falcioni L, et al., 2016. Paired serum and urine concentrations of biomarkers of diethyl phthalate, methyl paraben, and triclosan in rats. Environ. Health Perspect 124, 39–45. [PubMed: 26047088]
- US EPA, 1988. Recommendations for and documentation of biological values for use in risk assessment US Environmental Protection Agency, Cincinnati, OH.
- US EPA, 1991. Guidelines for developmental toxicity risk assessment EPA/600/FR-91/001. Washington DC.
- US EPA, 1996. Guidelines for reproductive toxicity risk assessment EPA/630/R-96/009. Washington, DC.
- US EPA, 2006. A framework for assessing health risk of environmental exposures to children EPA/600/R-05/093F. Washington, DC: Environmental Protection Agency.

- Williams GM, Iatropoulos MJ, 2002. Alteration of liver cell function and proliferation: Differentiation between adaptation and toxicity. Toxicol. Pathol 30, 41–53. [PubMed: 11890475]
- World Health Organization, 2003. Concise international chemical assessment document 52: Diethyl phthalate. RISKLINE/2003120009 Geneva.
- Wormuth M, Scheringer M, Vollenweider M, Hungerbuhler K, 2006. What are the sources of exposure to eight frequently used phthalic acid esters in Europeans? Risk Anal 26, 803–824. [PubMed: 16834635]
- Yost EE, Euling SY, Weaver JA, Beverly BEJ, Keshava N, Mudipalli A, et al., 2019. Hazards of diisobutyl phthalate (DIBP) exposure: A systematic review of animal toxicology studies. Environ. Int 125, 579–594. [PubMed: 30591249]
- Zota AR, Calafat AM, Woodruff TJ, 2014. Temporal trends in phthalate exposures: Findings from the National Health and Nutrition Examination Survey, 2001–2010. Environ. Health Perspect 122, 235–241. [PubMed: 24425099]



Fig. 1.

(A) Study evaluation criteria and (B) strength of evidence characterization for DIBP animal toxicology studies.



Fig. 2.

Literature flow diagram (based on PRISMA flow diagram) for identifying DEP animal toxicology studies. ^aOther sources consisted of forward and backward searches, searching citations from key references, manual search of citations from key regulatory documents, and references that had been previously identified from an earlier DEP review effort. ^bIncludes records identified from TSCATS2, TSCATS1 (searched via Toxline), and TSCA section 8e recent notices identified via Google search, as described in the protocol. ^cIncludes 4 supplementary materials (not main text articles) that were tagged as records during the literature search. These supplementary materials were not included in the count of records for title/abstract screening. ^dMost studies reported data on multiple hazards; see Table 1.

Weaver et al.

Table 1

List of studies and overall study confidence by $outcome.^a$

				Male reproduc	$b_{ m stive}$					Female reprodu	$c_{\rm ctive}c$						$d_{\text{ev.}}d$		I-I-M	2			Kidney		
Author (year)	Species (strain)	Exposure life stage and duration	Exposure route	Testoster one	Morphological development	Sperm	Fertility	Organ weight	Histopathology	Pregnancy outcomes	Maternal body weight	Organ Hi weight	stopathology F	Iormones E	strous Mo yclicity dev	rphological S elopment	urvival Gro	wth Struc alter	tural Org tions wei	an Histopathok ht	gy Biochemistry	Tumors	Organ weight	Histopathology	Biochemistry
Fujii et al. (2005)	Rat [Crj:CD (SD)]	Multigenerational study: F0 and F1 each exposed for 10 wests + mating, gestation, and weaning (~15-17 weeks total)	Diet	Ξ	Ξ	н	н	н	W	н	M	н		ш.	н	<u></u>	H	I	ж	Μ	Ξ	1	ж	W	1
RTI International (1984), Lamb et al. (1987)	Mouse (CD-1)	Multigenerational study (continuous breeding protocol)	Diet	I.	ı	н	Ξ	Ξ		н	I	T	I	1	I	-	н	I	н	I	I	I	I		I
Pereira and Rao (2007)	Rat (Wistar)	Multigenerational study; F0 and F1	Diet	I	I	I	I	I	I	I	I	1	I		I	I	Г	I	Г	Г	Г	I	I	I	I
Pereira et al. (2007c)		mating, gestation, and weaning (150)		i	I	I	I	I	I	I	I	I	I		I	I	1	I	Г	Г	Г	I	I	I	I
Pereira et al. (2007d)		days total)]; F2 exposed for 150 days after			I	I	I	г	I	I	I	L	I		I	I		I	I	I	I	I	ī	I	I
Pereira et al. (2007a)		weaning		I	I	I	I	I	I	I	I	I	I	1	I	I	1	I	Г	Г	Г	I	I	I	I
Hardin et al. (1987), Hazleton Laboratories America Inc (1983a, 1983b, 1922)	Mouse (CD-1)	GD 6–13	Gavage	I	ī	1	I	I	I	I	¥	1	I		1	-	ж 	I	I	I	I	I	I	I	I
NTP (1988), Field et al. (1993)	Rat (CD)	GD 6–15	Diet	I	I	I	I	I	I	I	н	Н	I	I	I	т	н	н	I	I	I	I	I	I	1
Procter & Gamble (1994)	Rabbit (New Zealand White)	GD 6–18	Dermal	I	I	I	I	I	I	1	L	н	I		I	T	н	н	I	I	I	I	I	I	I
Howdeshell et al. (2008)	Rat (Sprague -Dawley)	GD 8–18	Gavage	н	I	I	ī	I	I	I	M	I	I		I	-		I	I	I	I	I	I	I	I
Liu et al. (2005)	Rat (Sprague -Dawley)	GD 12–19	Gavage	1	н	I	I	I	I	I	I	I	I		I	I	I .	I	I	I	I	I	I	1	I
Furr et al. (2014)	Rat (Sprague -Dawley)	GD 14–18	Gavage	н	I	I	I	I	I	I	г	1	I		I	~	- V	I	I	I	ı	I	I	I	I
Gray et al. (2000)	Rat (Sprague -Dawley)	GD 14 - PND 3	Gavage	W	н	I	I	Н	I	I	M	I	I	I	Н	2	н	I	н	I	I	I	Н	I	I
Setti Ahmed et al. (2018)	Rat (Wistar)	GD 8 – PND 30	Gavage	I	I	I	I	I	I	I	I	1	I		I	Т	L ,	Г	Г	I	I	I	I	I	I
Manservisi et al. (2015)	Rat (Sprague -Dawlev)	PND 1–181 (females)	Gavage	I	I	I	I	I	I	Ц	I	Г -	I		I	I		I	I	I	I	I	I	I	I
Hu et al. (2016)				I	I	I	ı	I	I	I	1	1	I		I	I	Г	I	I	I	I	I	ı	I	I
Mapuskar et al. (2007)	Mouse (Swiss)	2–3 weeks old (females); 90- day exposure	Diet	I	I	I	I	I	I	I	I	1	I	1	I	I	1	I	Г	Ч	Г	I	I	I	I

uscript EPA Author Manuscript

EPA Author Manuscript

Dev. d Twee	Histopathology Hormones Extrons Morphological Survival Growth Structural Organ cyclicity development		Ξ Ι Ι Ι Ι Ι Ι	H I I I I I I	M 1 1 1	E 1 1				I				N 1 1 1 1 1 1		м н Г		
Female reproductive c	an Histopathology Pregnancy Maternal Organ ght outcomes body weght weght					1				1				1		H W	- - -	
Male reproductive b	Exposure Testosterone Morphological Sperm Fertility Org. route development weig	Diet L – – – M	Gauge H - L	Gavage L	Diet L	Demai	Diet	Drinking	Diet	Diet	Diet L – – – L	Diet	Gavage	Diet	Diet	Gavage H - M - L	Diet M - M	Drinking
	r Species Exposure life (strain) stage and duration	and Rat 5 weeks old a (Wistar) (males); 7-day) exposure	c et al. Rat 5 weeks old * (Sprague (males): 28-day -Dawley) exposure	et al. Rat 5 weeks old * (Sprague (males): 14-day -Dawley) exposure	et al. Rat (CD) Age nor reported, buttach yre- pubrertal, based on buttach yreight (malks and formal 2.42-day and 112-day exposures	1995) Mouse 6 weeks old (B6C3 (males and F1), Rats femlar): 28 day (F344A) or 104–105- week	t et al. Rat 6–7 weeks old (Wistar) (males): 150- day exposure	et al. Rat 7 weeks old (Sprague- (males); 120-day Dawley) exposure	and Rat 6–7 weeks old (006b) (Wistar) (females); 150- day exposure	and Rat 7–8 weeks old 006a) (Wistar) (males and females): 150-day exposure	et al. Rat 7-8 weeks old (Wistar) (males): 150- day exposure	et al. Rat 7-8 weeks old) (Wistar) (males): 150- day exposure	t al. Rat 6-8 weeks old (Wistar) (males); 2-day exposure	and Rat Age not reported (F334) (males); 21- day exposure	and Rat Age not reported (F334) (males); 21- day exposure	hi et Rat 8 weeks old 60 (Sprague- (males and Dawley females): 28-day exposure	l et Mice 8 weeks old 9) (Swiss (males): 90-day abino) exposure	and Rat 12 weeks old 007) (Wistar) (males and females); 180-day exposure

Weaver et al.

b Male reproductive endpoints: Testosterone (testicular production or level measured in testis or serum), morphological development (AGD, nipple retention, time to puberty), sperm (sperm counts, morphology), fertility (copulation and ability to produce offspring), organ weights (testis, epididymides, prostate, seminal vesicles), histopathology (general histopathological evaluations of male reproductive organs).

C reproductive organs), hormones (any measurement of reproductive hormones), estrous cyclicity, morphological development (AGD, time to puberty).

d Developmental endpoints: Survival (fetal viability, fetal mortality, resorptions, pre- or post-implantation loss, postnatal survival), growth (pre- or postnatal body weight), structural alterations (external, skeletal, or visceral malformations or variations).

EPA
Autho
r Manu
Jscript

Table 2

Evidence profile table for male reproductive effects of DEP or MEP.

<u>Male reproductive efi</u>	fects					
Outcome		Available Studies	Factors that increase confidence	Factors that decrease confidence	Confidence judgement for outcome	Confidence judgement for overall hazard
Gestational or postnatal exposure (including F1 and F2 offspring from multigenerational studies	Testosterone	High Confidence: Howdeshell et al. (2008), Furr et al. (2014) Medium Confidence: Gray et al. (2000)	 Consistency Lack of effect in sensitive species (Tat) Minimal concern for bias and sensitivity 	• Few studies	OOO INDETERMINATE There were no effects on fetal testosterone production at GD 18 (Howdeshell et al. 2008, Furr et al. 2014) or serum testosterone levels at PND 3 (Gray et al. 2000) following gestational exposure.	⊕⊕⊖ MODERATE Based on moderate evidence for adverse effects on sperm and decreased decreased textosterone in testosterone in
	Male morphological development	High Confidence: Fujii et al. (2005), Gray et al. (2000), Liu et al. (2005)	 Consistency Minimal concern for bias and sensitivity 	• Few studies	COO INDETERMINATE There were no effects on nipple retention or AGD (Gray et al. 2000, Liu et al. 2005) or on the age of preputial separation (Fujii et al. 2005, Gray et al. 2000) following gestational exposure.	
	Reproductive organ weights	High Confidence: Fujii et al. (2005), Gray et al. (2000)	 Consistency Minimal concern for bias and sensitivity 	• Few studies	NDETERMINATE There were no effects on male reproductive organ weights following gestational exposure, except for a decrease in absolute prostate weight and increase in relative seminal vesicle weight in F1 weanlings (Fujii et al. 2005).	
Peripubertal or adult exposure (including F0 and F1 parental animals from multigenerational studies)	Testosterone	High Confidence: Fujii et al. (2005). Shiraishi et al. (2006) Medium Confidence: Mondal et al. (2019) Low Confidence: Oishi and Hiraga (1980), Pereira et al. (2008a)	• Consistency in the direction of effect across three studies	 Low precision in the high confidence study by Fujii et al. 2005 Concerns for bias and sensitivity Mechanistic evidence indicating lack of effect on steroidogenesis 	⊕⊖⊖ SLIGHT Significantly decreased testosterone was observed in rats exposed as adults/young adults in the studies by Fujii et al. 2005, Pereira et al. (2008a), and Oishi and Hiraga 1980, although Fujii et al. 2005 tested a relatively small sample size and had high variability, and the latter two studies are considered <i>low confidence</i> for this outcome. No effects on testosterone were observed in adult male rats for a 28-day exposure (Shiraishi et	

Environ Int. Author manuscript; available in PMC 2021 December 01.

EPA
Autho
r Manu
uscript

Male reproductive effects

Weaver et al.

Outcome		Available Studies	Factors that increase confidence	Factors that decrease confidence	Confidence judgement for outcome	Confidence judgement for overall hazard
					al. 2006) or in mice after a 3-month exposure (Mondal et al. 2019).	
	Sperm	High Confidence: Fujii et al. (2005), RTI International (1984), Kwack et al. (2009) Medium Confidence: Shiraishi et al. (2019), Mondal et al. (2019)	 Minimal concern for bias and sensitivity Biological plausibility (mechanistic evidence of oxidative stress in gern cells) 	• Unexplained inconsistency	⊕⊕⊖ MODERATE MODERATE Decreased sperm number, increased incidence of abnormal sperm, and/or decreased sperm motility were observed in multiple studies that exposed adult rats or mice (Fujii et al. 2005, Kwack et al. 2019), although effects differed somewhat between studies and were not consistently observed. Shirashi et al. 2006 used a similar study design as Kwack et al. 2009 but did not observe any effects.	
	Fertility	High Confidence: Fujii et al. (2005), RTI International (1984)	 Consistency Minimal concern for bias and sensitivity 	• Few studies	OOO INDETERMINATE No effects on fertility in mating pairs were observed in two multigenerational reproductive studies.	
	Reproductive organ weights	High Confidence: RTI International (1984), Fujii et al. (2005) Medium Confidence: Oishi and Hiraga (1980) Low Confidence: Brown et al. (1978), Kwack et al. (2009, 2010), Pereira et al. (2007d, 2008a), Shiraishi et al. (2006)		 Unexplained inconsistency inconsistency and sensitivity and sensitivity 	NDETERMINATE Absolute Absolute apsolute (testes, epididymides, prostate, seminal vesicles) were generally not affected in <i>high</i> or <i>medium</i> <i>confidence</i> studies in mice and rats. Effects on relative testis weight were inconsistent across the remaining <i>low</i> <i>confidence</i> studies, although this is considered a less reliable measurement compared to absolute testis weight.	
	Histopathology	Medium Confidence: Fujii et al. (2005), Shiraishi et al. (2006), Jones et al. (1993) Low Confidence: Mondal et al. (1978) Brown et al. (1978)	• Coherent with effects on sperm	 Unexplained inconsistency Concerns for bias and sensitivity and sensitivity Quantitative data not reported in any study 	OOO INDETERMINATE Altered testicular histoarchitecture (Mondal et al. 2019) and altered Leydig cell cytoplasmic ultrastructure Jones et al., 1993) were (Jones et al., 1993) were reported, but no effects on histopathology of male repoductive organs was observed in three other studies.	

Π
Þ
Au
iho
r Z
lar
snı
crij
ę

Weaver et al.

Table 3

Evidence profile table for female reproductive effects of DEP.

Female reproducti	ve effects				
Outcome	Available studies	Factors that increase confidence	Factors that decrease confidence	Confidence judgement for outcome	Confidence judgement for overall hazard
Pregnancy outcomes	High Confidence: Fujii et al. (2005), RTI International (1984) Low Confidence: Manservisi et al. (2015)	Minimal concern for bias and sensitivity in most studies	 Few studies Small magnitude of effect in high confidence studies 	⊕⊖⊖ SLIGHT Statistically significantly decreased gestation length (Fujii et al. 2005) and decreased litter size (RTI International 1984) were observed in F1 parental females in two <i>high</i> confidence studies, and increased litter size was observed in one <i>low</i> confidence study (Manservisi et al. 2015). Otherwise, no effects were observed.	⊕⊖⊖ SLIGHT Based on few reported effects across studies, including effects on gestation length, litter size, and organ weights or histopathology.
Maternal body weight	High Confidence: NTP (1988) Medium Confidence: Howdeshell et al. (2008), Hardin et al. (19205), Gray et al. (2000), Fuji et al. (2005) Low Confidence: Furr et al. (2014), Procter & Gamble (1994)	 Minimal concerns for bias and sensitivity in two studies that observed effects 	• Unexplained inconsistency	OOO INDETERMINATE NTP (1988) reported a trend towards decreased maternal body weight gain (corrected for gravid uterine weight), whereas Fujii et al. (2005) reported increased maternal body weight gain in F0 females. Otherwise, no increased maternal body weight gain were observed in rats, mice, or rabbits.	
Reproductive organ weights	High Confidence: Fujii et al. (2005), RTI International (1984), Procter & Gamble (1994), NTP (1988), Shiraishi et al. (2006) Low Confidence: Brown et al., 1978; Pereira et al., 2007d		 Concems for bias and sensitivity in one out of two studies that observed effects 	⊕⊖⊖ SLIGHT Decreased absolue and relative uterine weights was observed in F1 and F2 absolue and relative uterine weights was observed in F1 and F2 ratis in a <i>high confidence</i> two-generation study (Fujii et al. 2005), although the effect was not observed in adult F1 animals so may be transient. A significant decrease in relative ovary weights was reported in a significant decrease in relative ovary weights was reported in a bw confidence study in F0 and F1 adult rats (Pereira et al., 2007c). Otherwise.	
Histopathology	High Confidence: Procter & Gamble (1994) Medium Confidence: Fujii et al. (2005), Shiraishi et al. (2006) Low Confidence: Manservisi et al. (2015)		• Concerns for bias and sensitivity in the only study that observed effects	⊕⊖⊖ SLIGHT A reduction in the size of the lobular structures of the mammary gland was observed in a <i>low confidence</i> study in parous F1 female rats (Manservisi et al. 2015). Otherwise, on histopathological changes were observed in ovaries, uteri, vaginas, or mammary glands.	

111
75
$\mathbf{\Sigma}$
~
≒
÷
ō
¥
-
_
<
Š
Mar
Manu
Manu
Manus
Manusc
Manuscri
Manuscrip

Female reproductiv	re effects				
Outcome	Available studies	Factors that increase confidence	Factors that decrease confidence	Confidence judgement for outcome	Confidence judgement for overall hazard
Hormones	High Confidence: Shiraishi et al. (2006)	Minimal concern for bias and sensitivity	Single study	OOO INDETERMINATE No effects were observed on steroid hormone or gonadotropin levels in female rats after subchronic exposure (Shiraishi et al. 2006).	
Estrous cyclicity	High Confidence: Fujii et al. (2005) Low Confidence: Shiraishi et al. (2006)	 Minimal concern for bias and sensitivity 	• Few studies	OOO INDETERMINATE No effects were observed on estrous cyclicity in F0 or F1 females in a two-generation study (Fujii et al. 2005) or after subchronic exposure (Shiraishi et al. 2006).	
Female morphological development	High Confidence: Fujii et al. 2005, Gray et al. 2000	 Minimal concern for bias and sensitivity 	Single study	OOO INDETERMINATE A statistically significant delay in the age at vaginal opening was observed in F1 rats in a two-generation study, likely related to decreased growth. No DEP- related effects on AGD were observed in F1 or F2 animals in this study (Fujii et al. 2005)	

EPA	
Authc	
or Mar	
nuscr	
ipt	

Table 4

EPA Author Manuscript

EPA Author Manuscript

Evidence profile table for developmental effects of DEP.

Developmenta	l effects				
Outcome	Available studies	Factors that increase confidence	Factors that decrease confidence	Confidence judgement for outcome	Confidence judgement for overall hazard
Fetal survival	High Confidence: Fujii et al. (2005) Procter and Gamble (1994) Hardin et al. (1987), Howdeshell et al. (2008), RTI International (1984), NTP (1988) Medium confidence: Fur et al. (2014), Gray et al. (2000) Fur et al. (2014) Setti Ahmed et al. (2018)	 Minimal concerns for bias and sensitivity in most studies 	• Effects observed in only one study	OOO INDETERMINATE A continuous breeding study reported decreased live F2 pups at birth (RTI International 1984). No other studies reported dose-related effects on fetal survival.	⊕⊕⊖ MODERATE Based primarily on the consistent reduction in postnatal growth observed across gestational and early postnatal exposure studies.
Postnatal survival	High Confidence: Fujii, et al. (2005), Hardin et al. (1987) Low Confidence: Manservisi et al. (2015)	 Minimal concerns for bias and sensitivity in most studies 	• Few studies • Unexplained inconsistency	OOO INDETERMINATE Decreased postnatal survival of F2 pups was observed in the <i>low confidence</i> study by Manservisi et al. (2015), and there was a nonsignificant trend towards decreased survival of F1 but not F2 pups in the high <i>confidence</i> study by Fujii et al. (2005). Otherwise, no effects were observed.	
Fetal growth	High Confidence Fujii et al. (2005), RTI International (1984), NTP (1988), Hardin et al. (1987), Procter & Gamble (1994), Gray et al. (2000) Low Confidence : Setti Ahmed et al. (2018)	 Consistency Minimal concerns for bias and sensitivity 		OOO INDETERMINATE All studies that assessed prenatal growth observed that DEP had little to no effect on fetal body weight or of weight at birth in rats, mice or rabbits.	
Postnatal growth	High Confidence Fujii et al. (2005), RTI International (1984), Hardin et al. (1987), Gray et al. (2000) Low Confidence: Pereira and Rao (2007), Manservisi et al. (2015), Oishi and Hiraga (1980), Hu et al. (2016), Setti Ahmed et al. (2018)	• Consistency across studies with longer exposure durations	• Concerns for risk of bias and sensitivity in some studies	⊕⊕⊕ ROBUST Decreased postnatal body weights were observed in offspring in the <i>high confidence</i> multigenerational exposure studies by Fujii et al. (2005) and RTI International (1984), including effects in both F1 and F2 generations in the study by Fujii et al. (2005), and in F1 animals in the <i>low confidence</i> studies by Pereira and Rao (2007), Maservisi et al. (2015), Hu et al. (2016), and Seti Ahmed et al. (2018). No effects on growth were observed following shorter duration gestational exposures (Hardin et al. 1987, Gray et al. 2000) or Periphertal exposure (Oishi and Hiraga 1980).	

EPA	
Author	
Manusc	
ript	

Development	al effects				
Outcome	Available studies	Factors that increase confidence	Factors that decrease confidence	Confidence judgement for outcome	Confidence judgement for overall hazard
Fetal structural alterations	High Confidence: NTP (1988), Procter & Gamble (1994) Low Confidence: Setti Ahmed et al. (2018)	• Exposure-response gradient in the study by NTP 1988 Minimal concern for bias and sensitivity	• Few studies	⊕⊖⊖ SLIGHT A dose-related increase in the percentage of fetuses with a rudimentary or supernumerary rib (a skeletal variation) was observed in rats exposed during gestation (NTP 1988). Findings in rabits (Proter & Gamble 1994) including fused ribs, missing lumbar, coccygeal vertebrae, hemia umbilicalis, and incurved ribs were not considered to be related to dose. Effects on intestinal development were observed in rats (Setti Ahmed et al. 2018).	

EPA
Author
Manus
cript

Weaver et al.

Table 5

Evidence profile table for liver effects of DEP or MEP.

Liver effects					
Outcome	Available studies	Factors that increase confidence	Factors that decrease confidence	Confidence judgement for outcome	Confidence judgement for overall hazard
Organ weight	High Confidence: Fujii et al. (2005), Gray et al. (2000), NTP (1995), RTI International (1984), Kwack et al. (2009), Kwack et al. (2010), Shiraishi et al. (2006), Shiraishi et al. (2006) Medium Confidence Hiraga (1980), Brown et al. (1978) Low Confidence: Mapuskar et al. (2007), Pereira et al. (2006, 2007c, 2007d, 2008b), Pereira and Rao (2006a, 2007), Setti Ahmed et al. (2018), Sonde et al. (2000)	• Consistency across most higher dose studies plausibility	 Inconsistent findings in low dose studies Low magnitude of effect 	⊕⊕⊖ MODERATE Dose-related increases in relative liver weight were observed in most studies that used higher dose levels of DEP, with statistically significant changes only observed at the highest doses tested (Fujii et al. 2005, Brown et al. 1978, NTP 1995, Moody and Reddy 1978, RTI International 1984, Otishi and Hirraga 1980). However, no effects on liver weight were observed in four <i>high confidence</i> rat studies (Gray et al. 2000, Kwack et al. 2009, 2010; Shiraishi et al. 2006) and effects on liver weight were effects on liver weight were inconsistent across <i>low confidence</i> studies.	⊕⊕⊖ MODERATE Based on evidence of increased liver weight, histopathological effects, and biochemical changes that are indicative of hepatic effects. However, effects. However, effects. However, evidence for histopathological and biochemical effects was primarily found in low does studies that had concerns for bias and sensitivity.
Histopathology	High Confidence: NTP (1995) Redium confidence: Fujir et al. (2005), Shiraishi et al. (2006), Moody and Reddy (1978) Moody and Reddy (1978) Low confidence: Brown et al. (1978), Mapuskar et al. (2007), Pereira et al. (2006, 2007a, 2007c, Sinkar and Rao (2006a, 2008b), Sinkar and Rao (2007)	• Biological plausibility	 Unexplained inconsistency inconsistency are generally not provided Concerns for bias and sensitivity in low dose studies that observed effects 	OOO INDETERMINAT Reports of intracellular vacuolations, degenerative changes in centrilobular and periportal areas, and necrosis were observed in <i>low confidence</i> low dose studies ranging from 0.57 mg/kg-day to 6.25 mg/kg- day, although these studies did not provide quantitative data. Mild or no these studies did not provide quantitative data. Mild or no liver histopathology were observed in studies that used higher dose levels.E	
Biochemistry	High Confidence: Fujii et al. (2005), NTP (1995), Kwack et al. (2009), Kwack et al. (2010) Medium Confidence: Moody and Reddy (1978), Moody and Reddy (1982), Shiraishi et al. (2006) Low confidence: Mapuskar et al., 2007; Pereira and Rao, 2007a, 2006b, 2008b; Sinkar and Rao, 2007; Sonde et al., 2000	• Biological plausibility	 Unexplained inconsistency Concerns for bias and sensitivity in low dose studies that observed effects 	⊕⊖⊖ SLIGHT <i>High confidence</i> sudies observed limited evidence of effects on biochemical markers of markers of liver damage, whereas <i>low confidence</i> low dose studies (0.57-6.25 mg/kg-day) reported dose-related increases in serum enzyme levels (e.g. ALT, AST, SOD) in both sexes of rats and mice across generations following oral exposures of at least 120 days. Alterations in triglycerides, glucose, cholesterol, and liver glycogen levels were	

	Confidence judgement for overall hazard	5)
	Confidence judgement for outcome	also reported primarily in the <i>low confidence</i> studies.
	Factors that decrease confidence	
	Factors that increase confidence	
	Available studies	
Liver effects	Outcome	

Evidence prof	ile table for kidney effects of DEP or	r MEP.			
Kidney effects					
Outcome	Available studies	Factors that increase confidence	Factors that decrease confidence	Confidence judgement for outcome	Confidence judgement for overall hazard
	High Confidence: Fujii et al. (2005), Kwack et al. (2009), Kwack et al. (2010), NTP (1995), Shiraishi et al. (2006), Gray et al. (2000) Medium Confidence: Oishi and Hiraga (1980), Brown et al. (1978)	 Minimal concern for bias and sensitivity in most studies 	• Unexplained inconsistency	OOO INDETERMINATE Inconsistent findings were reported for changes in kidney weight. Increases, decreases, or no effect in relative and absolute kidney weight were reported following oral or dermal DEP exposures in rats or mice, with no clear pattern of effect.	OOO INDETERMINATE Based on inconsistent effects on kidney weight across studies, limited biochemical findings, and lack of histopathological findings.
Organ weight	Low Confidence: Setti Ahmed et al. (2018)				
Histopathology	High Confidence: NTP (1995) Medium Confidence: Fujii et al. (2005), Shiraishi et al. (2006) Low Confidence: Brown et al. (1978)	 Minimal concern for bias and sensitivity in most studies Consistency 	• Few studies	OOO INDETERMINATE No treatment-related histopathological lesions in the kidney were reported in rats or mice following DEP exposure.	
Biochemistry	High Confidence Kwack et al. (2009), Kwack et al. (2010) Medium Confidence: Brown et al. (1978)	Minimal concern for bias and sensitivity	• Few studies	OOO INDETERMINATE Serum calcium levels were found to be increased in rats, with no effect on other biochemical parameters related to kidney injury (Kwack et al., 2009, 2010). Findings for urinary cell excretion were inconsistent (Brown et al. 1978).	

Table 6

EPA Author Manuscript

Cancer DutcomeAvailable studies increase confidenceFactors that decrease confidenceConfidence judgement for outcomeConfidence judgement for outcomeDutcomeBiologicalBiologicalSingle studyOOOOOOTumorsHigh Confidence NTPBiologicalSingle studyOOOOOOI 995PlausibilityNDETERMINATEOOOOOOI end mice following a 2-year bioasasy. A dose-related trend was observed in adenomas and carcinomas males; however, the increase was within the 2-year historical norms. In females, no dose response was observed.OOO	Evidence pro	ofile table for cancer.				
Tumors High Confidence NTP Biological Single study OOO 1995 plausibility INDETERMINATE INDETERMINATE 1995 plausibility Increased incidence of liver neoplasms was reported in male and female mice following a 2-year bioassay. A dose-related trend was observed in adenomas and carcinomas males; however, the increase was within the 2-year historical norms. In females, no dose response was observed.	Cancer Outcome	Available studies	Factors that increase confidence	Factors that decrease confidence	Confidence judgement for outcome	Confidence judgementfor overall hazard
	Tumors	High Confidence NTP 1995	Biological plausibility	Single study	OOO INDETERMINATE Increased incidence of liver neoplasms was reported in male and female mice following a 2-year bioassay. A dose-related trend was observed in adenomas and carcinomas males; however, the increase was within the 2-year historical norms. In females, no dose response was observed.	000 INDETERMINATE

EPA Author Manuscript

EPA Author Manuscript

Table 7