



EPA Public Access

Author manuscript

Environ Int. Author manuscript; available in PMC 2021 November 17.

About author manuscripts

Submit a manuscript

Published in final edited form as:

Environ Int. 2019 April ; 125: 579–594. doi:10.1016/j.envint.2018.09.038.

Hazards of diisobutyl phthalate (DIBP) exposure: A systematic review of animal toxicology studies

Erin E. Yost^{a,*}, Susan Y. Euling^{b,1}, James A. Weaver^a, Brandiese E.J. Beverly^{a,2}, Nagalakshmi Keshava^b, Anuradha Mudipalli^a, Xabier Arzuaga^b, Todd Blessinger^b, Laura Dishaw^a, Andrew Hotchkiss^a, Susan L. Makris^b

^aU.S. Environmental Protection Agency, National Center for Environmental Assessment, Research Triangle Park, NC, United States of America

^bU.S. Environmental Protection Agency, National Center for Environmental Assessment, Washington, DC, United States of America

Abstract

Background: Biomonitoring studies indicate a trend towards increased human exposure to diisobutyl phthalate (DIBP), a replacement for dibutyl phthalate (DBP). Recent reviews have found DIBP to be a male reproductive toxicant, but have not evaluated other hazards of DIBP exposure.

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 DIBP or the primary metabolite, monoisobutyl phthalate (MIBP).

Methods: A literature search was conducted in four online scientific databases [PubMed, Web of Science, Toxline, and Toxic Substances Control Act Test Submissions 2.0 (TSCATS2)], and augmented by review of regulatory sources as well as forward and backward searches. Studies were identified for inclusion based on defined 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.

Results: Nineteen toxicological studies in rats or mice met the inclusion criteria. There was *robust* evidence that DIBP causes male reproductive toxicity. Male rats and mice exposed to

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

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

¹Present affiliation: U.S. Environmental Protection Agency, Office of Children's Health Protection, Washington, DC.

²Present affiliation: National Institute of Environmental Health Sciences, National Toxicology Program, Office of Health Assessment and Translation, Research Triangle Park, NC.

Disclaimer

The views expressed are those of the authors and do not necessarily represent the views or policies of the U.S. Environmental Protection Agency.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.envint.2018.09.038>.

DIBP during gestation had decreased testosterone and adverse effects on sperm or testicular histology, with additional phthalate syndrome effects observed in male rats. There was also evidence of androgen-dependent and -independent male reproductive effects in rats and mice following peripubertal or young adult exposure to DIBP or MIBP, but confidence was reduced because of concerns over risk of bias and sensitivity in the available studies. There was also *robust* evidence that DIBP causes developmental toxicity; specifically, increased post-implantation loss and decreased pre- and postnatal growth. For other hazards, evidence was limited by the small number of studies, experimental designs that were suboptimal for evaluating outcomes, and study evaluation concerns such as incomplete reporting of methods and results. There was *slight* evidence for female reproductive toxicity and effects on liver, and *indeterminate* evidence for effects on kidney and cancer.

Conclusion: Results support DIBP as a children's health concern and indicate that male reproductive and developmental toxicities are hazards of DIBP exposure, with some evidence for female reproductive and liver toxicity. Data gaps include the need for more studies on male reproductive effects following postnatal and adult exposure, and studies to characterize potential hormonal mechanisms in females.

1. Introduction

Diisobutyl phthalate (DIBP) is a member of the phthalate ester class of chemicals and is used as a plasticizer to provide flexibility and durability to a wide variety of industrial and consumer products, including paints, lacquers, printing ink, pulp and paper, carpet, concrete, nail polish, and cosmetics (HSDB, 2017). Because of its use in household products, people are exposed to DIBP via food and indoor environments (Wormuth et al., 2006). DIBP is absorbed via oral ingestion (Koch et al., 2012) and dermal exposure (Elsisi et al., 1989), and is rapidly hydrolyzed to its primary metabolite, monoisobutyl phthalate (MIBP). DIBP and MIBP are distributed systemically in blood (Elsisi et al., 1989; Strucinski et al., 2006), and there is evidence these chemicals can be transferred to human breast milk (Fromme et al., 2011; Latini et al., 2009) and cross the placental barrier (Wittassek et al., 2009).

Biomonitoring studies indicate that DIBP exposures have increased in recent years, possibly because of DIBP being used as a substitute for other phthalates such as dibutyl phthalate (DBP) (Wittassek et al., 2007; Zota et al., 2014). For instance, data from the National Health and Nutrition Examination Survey (NHANES) indicate that the detection frequency of MIBP in urine increased from 72% of the U.S. general population in 2001–2002 to 96% in 2009–2010. Over this period, urinary concentrations of MIBP increased monotonically, while the metabolites of DBP and several other phthalates decreased (Zota et al., 2014).

Because DIBP was historically used less compared to other phthalates, it has also been relatively less studied. However, recent reviews have characterized DIBP as a male reproductive toxicant (CPSC, 2014; NAS, 2017). Male reproductive toxicity is a common hazard among many phthalates and is generally the greatest concern associated with phthalate exposure. In rats, in utero exposure to phthalates during the critical window of male sexual differentiation produces a phenotype known as “phthalate syndrome”, which is characterized by underdevelopment of male reproductive organs, decreased anogenital

distance (AGD), female-like nipple retention, cryptorchidism, and germ cell toxicity (CPSC, 2014; Foster and Gray Jr., 2013; Lioy et al., 2015; National Research Council, 2008). These effects can be causally linked to decreased testicular production of androgens, which are integral to male sexual development; decreased insulin-like-3 (INSL3) hormone, which regulates transabdominal testicular descent; and disruption of seminiferous cord formation, Sertoli cells, and germ cell development via an unknown mode of action (MOA) that is independent of effects on androgen production (Johnson et al., 2012; Martino-Andrade and Chahoud, 2010; National Research Council, 2008).

Based on recommendations following a review by a Chronic Hazard Advisory Panel (CHAP) (CPSC, 2014), which evaluated the effects on children's health of phthalates and phthalate alternatives used in children's toys and child care articles, DIBP is one of eight phthalates that the U.S. Consumer Products Safety Commission (CPSC) has permanently banned from children's toys and care articles at any amount greater than 0.1% [16 CFR § 1307 (2017)]. The CHAP stated that, although DIBP is not widely used in children's toys, it shares a similar anti-androgenic MOA with other banned phthalates, and therefore may contribute to cumulative risks to children. Additionally, the National Academy of Sciences (NAS) recently conducted a systematic review on phthalates and male reproductive tract development with the purpose of evaluating the potential for low-dose toxicity. The NAS concluded that DIBP is a presumed human health hazard based on dose-related effects on testosterone (T) production in animal studies (NAS, 2017).

Systematic review methods have only recently been applied for the purposes of chemical risk assessment, but offer the advantages of being focused, objective, and transparent. The CPSC (2014) and NAS (2017) reviews of phthalates both focused on evaluating epidemiological and toxicological studies with male reproductive outcomes, of which the NAS focused specifically on three phthalate syndrome endpoints (T, anogenital distance, and hypospadias) after gestational exposure. These systematic reviews did not evaluate other hazards caused by phthalate exposure, such as female reproductive effects or effects in other organ systems.

To gain a more comprehensive understanding of the spectrum of effects after DIBP exposure, we performed a systematic review of the animal toxicology literature for DIBP for six broad hazard categories (male reproductive, female reproductive, developmental, liver, kidney, cancer) that have been commonly associated with phthalate exposure.

2. Methods

This systematic review is part of a larger evaluation of health effects after exposure to phthalates, which includes systematic reviews of epidemiological and animal toxicology studies for multiple phthalates that will be published as separate manuscripts. 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”).

2.1. Literature searches and screening

A literature search was conducted in four online scientific databases [PubMed, Web of Science, Toxline, and Toxic Substances Control Act Test Submissions (TSCATS2)], using search terms designed to capture all potentially pertinent studies. Initial database searches were conducted in February 2013, with updates performed every 6–12 months through July 2017 (see protocol Section 3). 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 DIBP review effort and added to EPA's Health and Environmental Research Online (HERO) database.

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 DIBP animal toxicology studies (see protocol Section 2 for the full PECO):

- Population: Nonhuman mammalian animal species (whole organism) of any life stage.
- Exposure: Any administered dose of DIBP or MIBP as singular compounds, via oral, dermal, or inhalation routes of exposure.
- Comparator: Exposure to vehicle-only or untreated control
- Outcome: Any examination of male reproductive, female reproductive, developmental, liver, kidney, or cancer outcomes.

Title/abstract and full text screening was 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 (https://hero.epa.gov/hero/index.cfm/project/page_id/2320).

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 exposures 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).

All study evaluation ratings are documented and publicly available in EPA's version of Health Assessment Workspace Collaborative (HAWC), a free and open source web-based software application (<https://hawcprd.epa.gov/assessment/497/>). 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 (see protocol Section 5). Information was extracted on the study design (species, strain, life stage, exposure characteristics), outcome measurements, and statistical significance as reported by the study authors. Dose levels are presented as mg/kg-day. For dietary exposure studies, 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, data were synthesized and evaluated across studies according to the age and developmental stage of exposure to account for life stage-specific windows of susceptibility, as recommended by the US EPA (2006). The following considerations were used to articulate the strengths and weaknesses of the available evidence for each outcome [adapted from Hill, 1965; see protocol Section 6.1]: consistency, biological gradient (dose-response), strength (effect magnitude) and precision, biological plausibility, and coherence. Data were synthesized using a narrative approach, with no meta-analysis performed due largely to the heterogeneity of endpoints and study designs considered in this review (see protocol Section 6.2). Syntheses were based primarily on studies of *High* and *Medium* confidence. *Low* confidence studies were generally used to evaluate consistency, or if no or

few higher confidence studies are available. 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 protocol Section 6.3), using the framework outlined 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. *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 there are no studies available for that evidence stream or the evidence is inconsistent and cannot provide a basis for making a conclusion in either direction. *Compelling evidence of no effect* represents a situation where extensive evidence across a range of populations and exposures identified no association.

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; Schunemann 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 1156 unique records for DIBP, of which 31 proceeded to full text screening. Of these, 25 met the defined PECO criteria for inclusion and were moved forward for study evaluation. The studies excluded after full text screening were all animal toxicology studies that did not meet PECO criteria, consisting of three studies that used intraperitoneal injection (Lawrence et al., 1975; Ray et al., 2012; Singh et al., 1972), two studies in nonmammalian species [chickens: Choy, 1975; zebrafish: Sohn et al., 2016], and one behavioral study in mice (Ma et al., 2013). Additionally, the included article by University of Rochester (1954) presented data from multiple experiments, including a rat study that complied with the PECO criteria and was moved forward for study evaluation, and a dog study and several acute studies that were rejected for not meeting the PECO criteria.

At the study evaluation phase, one article was excluded from further consideration because of critical deficiencies in reporting (Eastman Kodak, 1954), and four articles (all of which presented data from the same reproductive toxicity study in rats) were excluded because they tested a single high dose of DIBP (4000 mg/kg-day) that caused a high rate of mortality in the exposed dams (Hardin et al., 1987; Hazleton Laboratories, 1983a, 1983b, 1992). Two of the included articles (Foster et al., 1981; Foster et al., 1982) presented overlapping data from the same animals.³ Therefore, a total of 20 articles presenting data from 19 unique studies were included in this analysis (Table 1).

3.2. Summary of included studies

Table 1 summarizes the experimental designs and outcomes evaluated in the included studies. All were oral exposures (gavage or diet) of rats or mice. Ten were prenatal developmental toxicity studies that dosed pregnant dams with DIBP; of these, nine exposed rats during gestation only, and one exposed mice either during gestation only or from gestation through weaning (via lactational exposure). The remaining nine studies were postnatal-only exposures of weanling or peripubertal rats or mice to DIBP or MIBP.

Most of the gestational exposure studies were designed with the primary objective of evaluating male reproductive effects in F1 offspring, although two rat studies focused on fetal survival and structural alterations (external, skeletal, and soft tissue/visceral malformations and variations) (BASF, 2007; Saillenfait et al., 2006). Because of the focus on male reproductive effects, four of the rat studies exposed animals only during the critical window of male sexual differentiation (Furr et al., 2014; Hannas et al., 2011; Hannas et al., 2012; Saillenfait et al., 2017), which is approximately between gestation days (GDs) 14–18. Other studies exposed pregnant dams for longer durations, beginning immediately after mating (mice: Wang et al., 2017), within a few days after implantation (rats: BASF, 2007; Howdeshell et al., 2008; Saillenfait et al., 2006), or at mid-gestation (rats: Saillenfait et al., 2008), and continuing through the remainder of gestation. One study in rats (Saillenfait et al., 2008) and one in mice (Wang et al., 2017) allowed dams to give birth and evaluated male reproductive effects in adult F1 offspring, while the remaining rat studies sacrificed dams near the end of gestation and evaluated effects in fetuses. Regardless of the primary objective of the study, all gestational exposure studies provided relevant data on F1 offspring survival, growth, and/or maternal reproductive outcomes.

Six of the postnatal exposure studies exposed sexually immature male rats or mice of varying ages (weanling to peripubertal) for up to 7 days and focused on male reproductive outcomes (Foster et al., 1981; Foster et al., 1982; Oishi and Hiraga, 1980a, 1980b, 1980c, 1980d; Zhu et al., 2010). One study by Sedha et al. (2015) focused on estrogenic outcomes in weanling female rats. Two studies by the University of Rochester (1953, 1954) exposed male or female rats for 1 or 4 months beginning at weaning and focused on general toxicity. All reported information on postnatal growth, and most also reported information on liver and kidney effects.

None of the gestational or postnatal exposure studies provided information on cancer.

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 for gestational exposure studies (Fig. S1) and postnatal exposure studies (Fig. S2). Figs. S1 and S2 each provide links to HAWC, where rationale for the study evaluation ratings is documented.

³Here, we cite Foster et al., 1981 (a peer-reviewed publication) for testis weights and histopathology, and Foster et al., 1982 (an edited book) for liver and kidney weights from the same study.

Among the gestational exposure study outcomes, confidence was generally *high*, although some study outcomes were rated *medium* or *low confidence* because of specific concerns. For instance, the rat gestational exposure studies by Furr et al. (2014) and Hannas et al. (2011, 2012) used a relatively small sample size (generally n = 3–4 dams/treatment group) that authors stated did not provide the statistical power to consistently detect anything but rather large alterations in maternal weight gain and fetal viability, or decreases in T production of less than 20–25%; our reviewers determined that the T measurements in these studies had adequate sensitivity, but confidence in the maternal weight gain and fetal survival data was reduced because of sensitivity concerns. For evaluation of maternal toxicity, confidence was reduced in the maternal body weight gain measurements in several studies that did not adjust for gravid uterine weight, the calculation of which is considered preferable because it facilitates the interpretation of maternal toxicity relative to effects on fetal body weight (US EPA, 1991). Additionally, in the only available gestational exposure study in mice (Wang et al., 2017), confidence in male reproductive, body weight, and organ weight data was reduced because data were presented as an average of individual pups, rather than using the litter as the experimental unit. Failure to account for litter effects has the potential to overestimate the statistical significance of experimental findings (Haseman et al., 2001). All rat gestational exposure studies used the litter as the experimental unit for presentation and analysis of F1 offspring data.

All outcomes reported in the postnatal exposure studies were rated as *medium* or *low confidence*. In general, these studies had incomplete reporting of experimental designs and results, which led to concerns for risk of bias and sensitivity. Concerns included lack of information about the age or pubertal status of the animals at the time of exposure, strain of the animals, sample size, or methods used to allocate test animals to experimental groups, and presentation of only qualitative (rather than quantitative) results. For evaluation of male reproductive effects, confidence was reduced in studies that presented only relative testis weight (as a percentage of body weight) without presenting absolute testis weight; it has been shown that testis weights are not modeled well by an organ-to-body weight ratio because testis and body weights are not proportional (Bailey et al., 2004), so this may be a less sensitive measurement compared to absolute testis weight.

3.4. Male reproductive effects

Figures indicating the doses at which statistically significant male reproductive effects occurred are provided in Supplemental Materials (Figs. S3 - S5). Morphological and histopathological effects in F1 postnatal and adult males from the multi-dose study by Saillenfait et al. (2008) are summarized in Figs. S6 and S7, respectively.

3.4.1. Summary of gestational (F1) exposure studies—A dose-related decrease in androgens was observed in male offspring from all rat and mouse gestational exposure studies that evaluated this outcome. This includes six studies that evaluated fetal testicular T production or testicular T or androstenedione (AN) levels in fetal rats (Borch et al., 2006; Furr et al., 2014; Hannas et al., 2011; Hannas et al., 2012; Howdeshell et al., 2008; Saillenfait et al., 2017), and one study that evaluated serum and testicular T levels in adult mice that had been exposed during gestation and/or through weaning (Wang et al., 2017).

Across these studies, changes in T were statistically significant at doses as low as 200 mg/kg-day, and the magnitude of effect was often large (decreased up to 96% compared to controls). Several of these studies also reported decreased testicular expression of genes or proteins in the steroidogenesis pathway in rats (Borch et al., 2006; Hannas et al., 2011; Hannas et al., 2012; Saillenfait et al., 2017) and in mice (Wang et al., 2017), providing mechanistic evidence in both species to corroborate the observed decrease in fetal testicular testosterone. The affected genes or proteins included steroid acute regulatory protein (StAR) and scavenger receptor class B type 1 (SR-B1), which are involved in cholesterol uptake and transport; and 3 β -hydroxysteroid dehydrogenase (3 β -HSD), cytochrome P450 (CYP) side chain cleavage (P450scc or CYP11A1), and CYP17A1, which are steroidogenic enzymes. Hannas et al. (2012) also reported decreased testicular gene expression of INSL3 in rats. Overall, the evidence for decreased fetal testicular T production after gestational exposure was found to be *robust*.

In postnatal and adult male rats that had been exposed during the critical window of male sexual differentiation during gestation, Saillenfait et al. (2008) reported numerous dose-related outcomes that are consistent with decreased testosterone and INSL3: decreased AGD, increased time to puberty (preputial separation), nipple retention, increased rate of external malformations of the reproductive tract (hypospadias, exposed os penis, cryptorchidism,⁴ cleft prepuce; Fig. S6), and decreased reproductive organ weights (testis, epididymides, seminal vesicles, prostate). Other gestational exposure studies that evaluated fetal rats also observed decreased male AGD (Borch et al., 2006) and increased incidence of cryptorchidism, with approximately two-thirds of testes in the highest dose group (1000 mg/kg-day) located in the upper half of the abdominal cavity (Saillenfait et al., 2006). Conversely, in weanling and adult mice, male AGD was not affected and effects on testis weight were inconsistent (Wang et al., 2017). It is not clear whether this reflects a species sensitivity difference or a difference in study design. Overall, based on strength of the data in rats, the evidence for effects on male morphological development after gestational exposure was found to be *robust*, and evidence for effects on male reproductive organ weight was found to be *moderate*. In these same animals, histopathological analysis found dose-related increases in lesions in the fetal rat testis (Borch et al., 2006) and adult rat testis (Saillenfait et al., 2008), with oligospermia or total azoospermia⁵ in the corresponding epididymides (Saillenfait et al., 2008, Fig. S7). Adult male mice exposed to DIBP during gestation had decreased sperm concentration and motility, which reached statistical significance in mice that were exposed from gestation through weaning (Wang et al., 2017). These outcomes are consistent with effects on sperm development and function that have been observed for other phthalates, which are thought to be mediated via androgen-dependent and -independent pathways (Johnson et al., 2012; National Research Council, 2008). Given the exposure-response gradient and consistency across species, the evidence for effects on testicular histology and sperm parameters was found to be *robust*.

⁴Cryptorchidism is reported as “nonscrotal testis” in adult males in Saillenfait et al., 2008 and “ectopic testis” in male fetuses in Saillenfait et al., 2006.

⁵Azoospermia refers to a lack of spermatozoa, and oligospermia refers to reduced spermatozoa.

3.4.2. Summary of postnatal (weanling or peripubertal) exposure studies—

Effects on androgens (measured as testicular T or serum T or dihydrotestosterone [DHT] levels) were inconsistent across a series of studies by Oishi and Hiraga (1980a, 1980b, 1980c, 1980d), in which immature (~5-week-old) rats or mice were exposed to DIBP or MIBP in the diet for 7 days. In rats, androgen levels were increased compared to controls (Oishi and Hiraga, 1980b, 1980c), whereas testicular T levels in mice were similar (Oishi and Hiraga, 1980a) or significantly decreased (Oishi and Hiraga, 1980d) compared to controls. The life stage of these animals at the time of exposure was inferred to be peripubertal, based on the age of the animals; however, only one of these studies confirmed the pubertal status of the animals [Oishi and Hiraga, 1980c reported that the rats had positive sperm counts and were between puberty and sexual maturity, whereas the pubertal status of animals in the other three studies was not reported]. Susceptibility to phthalate-induced effects on androgen levels is dependent on the life stage of exposure, with sexually immature animals being more susceptible (Albert and Jégou, 2014; Hotchkiss et al., 2008; National Research Council, 2008). The discord in findings could therefore be based on issues such as timing of exposure and timing of assessment. Given these study concerns and the inconsistency in results across studies, the evidence for effects on T was found to be *indeterminate*.

Despite the inconsistent findings for androgen levels, effects on male reproductive organ weight provide support for an androgen-dependent MOA in animals exposed during weanling and/or peripubertal life stages. A dose-related decrease in absolute and relative testis weight was consistently observed across all rat studies (Foster et al., 1981; Oishi and Hiraga, 1980b, 1980c; University of Rochester, 1954; Zhu et al., 2010), with no significant effects on prostate or seminal vesicle weight (Foster et al., 1981). Decreased absolute testis weights were also observed in 21-day-old mice exposed to DIBP for 7 days (Zhu et al., 2010), although statistically significant effects occurred at higher doses compared to rats in the same study. Conversely, 5-week-old mice exposed to DIBP or MIBP for 7 days had a statistically significant increase in relative testis weight compared to control (Oishi and Hiraga, 1980a, 1980d). Relative testis weight has been found to be an unreliable metric because testis and body weights are not proportional (Bailey et al., 2004), so this result is considered *low confidence* since it could be an artifact of the decreased body weights in the DIBP- and MIBP-treated mice. Nevertheless, results reflect a potential difference in species sensitivity, with rats being more sensitive than mice for effects on testis weight. Overall, reduced testis weight was observed in both species and was consistent with effects observed in rats exposed to DIBP during gestation, so the evidence for effects on male reproductive organ weight was found to be *robust*.

DIBP- or MIBP-exposed rats had increased testicular atrophy (defined as marked or total loss of germinal epithelium of the seminiferous tubules) (Foster et al., 1981) and decreased spermatocytes and spermatogonia compared to controls (Foster et al., 1981; Oishi and Hiraga, 1980b). These results are considered *low confidence* because authors provided little information on their methods, and neither study provided quantitative results for the observed effects on sperm. However, results are consistent with those observed in rodents exposed to DIBP during gestation (Saillenfait et al., 2008; Wang et al., 2017), and

are corroborated by the reported decrease in testis weight in these studies. Evidence for testicular degeneration and effects on sperm was therefore found to be *moderate*.

3.4.3. Synthesis of results for male reproductive effects—Overall, the available studies in rodents provide *robust* evidence that DIBP causes male reproductive toxicity (Table 2). Across the gestational exposure studies, male rats had decreased fetal testicular T and displayed the hallmarks of phthalate syndrome (reduced male reproductive organ weights, hypospadias, cryptorchidism, nipple retention, reduced AGD, germ cell effects), and male mice had decreased T and effects on sperm. These effects are consistent with well-established MOAs for male reproductive effects of phthalates, which act through both androgen-dependent (decreased steroidogenesis leading to decreased androgens) and androgen-independent (decreased INSL3 and germ cell effects) pathways, and are supported by mechanistic data in both rats and mice showing decreased fetal testicular expression of steroidogenic genes/proteins and INSL3. While not all studies reported statistical significance, the direction of effect observed for most outcomes was consistent across the available gestational exposure studies, although some biomarkers observed in rats (decreased AGD and decreased adult male reproductive organ weights) were not observed in the mouse study by Wang et al. (2017).

The database of postnatal exposure studies had more risk of bias and sensitivity concerns compared to the gestational exposure database, but likewise provides evidence of both the androgen-dependent MOA (decreased testis weights) and androgen-independent MOA (increased testicular atrophy and decreased spermatocytes). Although androgens were inconsistently affected across the available postnatal exposure studies for DIBP, this may simply reflect limitations in the study designs, such as uncertainty about the pubertal status of the animals. Pubertal exposure studies for other antiandrogenic phthalates [e.g. DBP and diethylhexyl phthalate (DEHP)] report decreased T as well as androgen-dependent and -independent effects on reproductive development (Noriega et al., 2009; US EPA, 2007).

Strengths of this database include the availability of several *high confidence* gestational exposure studies, with one large multi-dose study that assessed multiple outcomes in postnatal and adult animals that had been exposed to DIBP in utero during the critical window for male reproductive development (Saillenfait et al., 2008). Other strengths are the availability of studies in both rats and mice that assessed similar outcomes, including multi-dose studies in multiple rat strains. Although none of the studies included a direct measurement of male fertility (e.g. mating success), the observed increase in hypospadias, testicular atrophy, and germ cell effects (including azoospermia) suggest that fertility can be affected.

3.5. Female reproductive effects

Figures indicating the doses at which statistically significant female reproductive effects occurred are provided in the Supplementary Materials (Figs. S8 - S9).

3.5.1. Summary of gestational (F1) and postnatal (weanling) exposure studies—Following gestational exposure, Saillenfait et al. (2006) observed a non-significant increase in the incidence of fetuses with displaced ovaries. Borch et al. (2006)

reported a significant increase in female AGD/cubic root of body weight, whereas Saillenfait et al. (2008) found a slight but not statistically significant increase in female AGD following gestational exposure with body weight used as a covariate in their statistical analysis. A separate publication of mechanistic data for the Borch et al. (2006) animals (Boberg et al., 2008) reported elevated ovarian aromatase gene expression in DIBP-exposed female fetuses, suggesting a possible endocrine mechanism; ovarian estradiol was also measured but was found to be near the detection limit of the assay with no statistically significant differences between groups (data not shown by authors). In a 20-day pubertal assay in rats, no vaginal opening had occurred in control or DIBP-treated animals when the study was terminated at PND 41, indicating that DIBP did not accelerate the timing of female puberty (Sedha et al., 2015). Although data is limited and the biological implications are unclear, the effects on AGD and ovarian displacement suggest morphological changes in females that are analogous to those observed in male offspring from these studies. Overall, the evidence for effects on female morphological development is considered *slight*.

In female weanling rats exposed to DIBP in a 3-day uterotrophic assay or 20-day pubertal assay, Sedha et al. (2015) found no effects on uterine, paired ovary, or vaginal weight. Because only one study was available, the evidence for effects on reproductive organ weight in developing females is considered *indeterminate*.

3.5.2. Summary of maternal (F0) exposure studies—Only one study (Saillenfait et al., 2008) evaluated gestation length, and reported that there was no effect of DIBP exposure. The evidence for effects on gestation length was considered *indeterminate*.

Maternal body weight parameters were reported in all gestational exposure studies. Of the studies judged to be *high confidence* for this outcome, BASF (2007) reported a significant decrease in maternal body weight gain in Wistar rats after correcting for gravid uterine weight, whereas three other rat studies reported no effect on corrected maternal body weight gain (Borch et al., 2006; Saillenfait et al., 2006; Saillenfait et al., 2017). The remaining studies were judged to have lower confidence because they did not correct for gravid uterine weight. Of these, Howdeshell et al. (2008) observed a decrease in maternal body weight gain, which appears to be a consequence of fetal toxicity (decreased fetal body weights); and the remaining studies reported no effect on maternal weight gain (Furr et al., 2014; Hannas et al., 2011; Hannas et al., 2012; Saillenfait et al., 2008; Wang et al., 2017). Overall, the evidence for effects on maternal body weight gain are considered *slight*.

Saillenfait et al. (2006) reported a significant decrease in gravid uterine weight, which appears to be a secondary effect to reduced fetal body weights in this study, whereas BASF (2007) and Saillenfait et al. (2017) reported no effect on gravid uterine weight. Given this limited amount of information, the evidence for effects on maternal organ weights is considered *indeterminate*.

3.5.3. Synthesis of results for female reproductive effects—Overall, the available studies provide *slight* evidence that DIBP causes female reproductive toxicity following developmental or maternal exposure (Table 3). This database has significant limitations. The exposure durations used in the gestational exposure studies were selected

primarily for evaluating male reproductive toxicity and fetal morphological development, and may not be the most sensitive for detecting maternal toxicity. None of the studies included pre-mating, mating, or lactational exposure intervals; therefore, effects on fertility, mating, fecundity (other than litter size effects due to post-implantation loss), and estrous cyclicity were not evaluated. Additionally, limited data were presented on effects in F1 females exposed during gestation, compared to the detailed evaluations of reproductive toxicity in F1 males.

3.6. Developmental effects

Figures indicating the doses at which statistically significant developmental effects occurred are provided in the Supplementary Materials (Figs. S10 - S12).

3.6.1. Summary of gestational (F1) exposure studies—Two *high confidence* studies (Howdeshell et al., 2008; Saillenfait et al., 2006) that exposed Sprague-Dawley rats by gavage beginning soon after the time of implantation (GD 8–18 and GD 6–20, respectively) observed a dose-related increase in fetal resorptions per litter, leading to a 52–62% decrease in the number of live fetuses per litter. In contrast, effects on fetal viability were not observed in the other studies that reported this outcome, including two dietary studies that exposed Wistar rats for similar durations (BASF, 2007; Borch et al., 2006), four studies that exposed Sprague-Dawley rats for shorter durations in mid- to late gestation (Furr et al., 2014; Hannas et al., 2011; Hannas et al., 2012; Saillenfait et al., 2008), or in mice exposed throughout gestation (Wang et al., 2017). This variability in results appears to be driven by differences in species, strain, and study design (particularly exposure duration). Reduced fetal survival has also been observed for other phthalates, such as DBP, and evidence suggests that DBP can interfere with pregnancy maintenance by decreasing ovarian progesterone production (Gray Jr et al., 2006). Given this biologically plausible mechanistic hypothesis for other phthalates, and the large effect size and exposure-response gradient observed in Howdeshell et al. (2008) and Saillenfait et al. (2006), the evidence for effects on fetal survival was found to be *robust*.

Postnatal survival following gestational exposure was only evaluated by Saillenfait et al. (2006), who found no effect on the survival of male and female rat pups through weaning. The evidence for effects on postnatal survival was found to be *indeterminate*.

Most studies in rats reported a dose-related decrease in fetal body weights (BASF, 2007; Borch et al., 2006; Saillenfait et al., 2006; Saillenfait et al., 2017) and postnatal and adult body weights (Saillenfait et al., 2008) of animals exposed to DIBP during gestation. No effects on fetal body weight were observed in rats by Saillenfait et al. (2017), and Wang et al. (2017) observed no effects in mice at PND 21 but increased body weight at PND 80, although these two studies tested a single lower dose of DIBP. Given the exposure-response gradient observed in studies that tested a wider range of DIBP doses, the evidence for effects on prenatal growth was found to be *robust*, and evidence for effects on postnatal growth after gestational exposure was found to be *moderate*.

Different results were observed in the two studies that performed detailed analyses of fetal external, skeletal, and soft tissue/visceral malformations and variations following

exposure from GD 6–20. In Sprague-Dawley rats, Saillenfait et al. (2006) reported a dose-related increase in the total incidence of external, skeletal, and visceral malformations in fetuses following gestational exposure to DIBP, reaching statistical significance at 750 and 1000 mg/kg-day. In particular, the incidence of fused sternebrae increased statistically significantly with dose, and a variety of other individual malformations of the neural tube, eye, vessels of the heart, vertebral column, and axial skeleton exhibited dose-related increases that were not significant. Dose-related skeletal variations included supernumerary ribs, incomplete ossification of thoracic or lumbar vertebral centra, and assorted variations of the skull and axial skeleton. Dose-related visceral variations included dilation or distention of the ureter and dilated renal pelvis, as well as ectopic testis and displaced ovaries (discussed in the male and female reproductive sections, respectively). Comparatively, in Wistar rats, BASF (2007) did not observe any dose-related trends at doses up to 942 mg/kg-day. In some cases, the frequency of skeletal variations was significantly elevated in treatment groups relative to control; however, the incidences of these variations were consistently within the range of historical control data for that laboratory, suggesting that they were not caused by DIBP treatment. Additionally, no effects on external morphology were observed by Saillenfait et al. (2017) at the single dose tested (250 mg/kg-day). Given that a clear exposure-response gradient was observed in one rat study while no effects were observed in the others, the evidence for effects on fetal structural alterations was found to be *slight*.

3.6.2. Summary of postnatal (weanling or peripubertal) exposure studies—A dose-related decrease in postnatal and adult growth was observed in all available studies that exposed immature rats and mice (Oishi and Hiraga, 1980a, 1980b, 1980c, 1980d; Sedha et al., 2015; University of Rochester, 1953, 1954), which were judged to be *medium* or *low confidence* for this outcome. One limitation is that the studies by Oishi and Hiraga reported decreased food consumption in DIBP and MIBP treated animals, suggesting a palatability issue, so the decreased growth in these animals may have been a secondary effect related to food intake. The studies by University of Rochester did not report food consumption, but also used high doses of DIBP in diet that may have caused palatability issues. Overall, the evidence for effects on growth after postnatal exposure was found to be *moderate*.

3.6.3. Synthesis of results for developmental effects—Overall, the available studies in rodents provide *robust* evidence that DIBP causes developmental toxicity (Table 4). The strongest evidence was available from the gestational exposure studies, which found effects on survival, growth, and fetal structural alterations. One limitation is that many of the studies exposed fetuses only during the major window of male sexual differentiation (e.g. GD 14–18), which may not be the most sensitive for detecting effects on survival or on skeleton or organ system development, and there were no two-generation or continuous breeding studies available. This limits the ability to fully evaluate potential effects on the developing fetus; for instance, none of the rat studies exposed dams prior to implantation, so it was not possible to evaluate effects on preimplantation loss. Nevertheless, the targeted dosing periods are informative of developmental hazards overall, and may provide a starting point for understanding the windows of susceptibility for various outcomes in the developing fetus.

3.7. Liver effects

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

3.7.1. Summary of available studies—Dose-related increases in relative liver weight were observed in all studies that reported this outcome, consisting of postnatal exposure studies in male rats (Foster et al., 1982; Oishi and Hiraga, 1980b; University of Rochester, 1953, 1954), female rats (University of Rochester, 1954), and male mice (Oishi and Hiraga, 1980a, 1980d); and in weanling and adult male mice that were exposed during gestation (Wang et al., 2017). While the University of Rochester studies did not include a statistical analysis, the average magnitude of effect was large (increased by up to 84% compared to controls); and in all other cases, effects on relative liver weight were statistically significant. Statistically significant increases in absolute liver weight were also observed in some cases (Oishi and Hiraga, 1980b; Wang et al., 2017). The largest increases in liver weight were often seen at doses associated with marked growth retardation, which has the potential to exaggerate the effect of relative weight measurements; however, since the relationship between liver weight and body weight is proportional, changes in relative liver weight are considered informative of toxicity (Bailey et al., 2004). Given the consistent direction of effect across studies and two species, the evidence for effects on liver weight is considered *robust*.

Two of these studies also included a histopathological analysis (University of Rochester, 1953, 1954), and found no difference between treatment and control livers. However, these results are considered *low confidence* because of reporting limitations in these studies. Therefore, the evidence for histopathological effects in the liver is considered *indeterminate*.

3.7.2. Synthesis of results for liver effects—Change in organ weight alone has been used as a potentially sensitive indicator for toxicity from chemical exposure (Bailey et al., 2004). In the absence of histopathological changes or supporting biochemical or mechanistic evidence (e.g. changes in hepatic enzyme expression), however, the biological significance of the reported changes in liver weight is inconclusive (Hall et al., 2012). Taken together, the evidence for liver toxicity is considered *slight* (Table S1).

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. S14).

3.8.1. Summary of available studies—Dose-related increases in relatively kidney weight were reported in the 1- and 4-month exposure studies in male and female rats by the University of Rochester (1953, 1954), with more pronounced changes in males compared to females (University of Rochester, 1954), although the authors did not perform a statistical analysis. In contrast, the 7-day exposure studies found that relative kidney weight was statistically significantly decreased in mice exposed to DIBP (Oishi and Hiraga, 1980a), and not significantly changed in mice exposed to MIBP (Oishi and Hiraga, 1980d) or rats exposed to DIBP (Oishi and Hiraga, 1980b). In a 4-day exposure of rats to MIBP, relative

kidney weight was increased by 12% compared to controls, although the effect was not significant (Foster et al., 1982). Given these inconsistencies and the limited number of studies, the overall evidence for effects on kidney weight is considered *slight*.

Two of these studies also included a histopathological analysis (University of Rochester, 1953, 1954), and found no difference between treatment and control livers. These histopathology results are considered *low confidence* because of reporting limitations in these studies. The evidence for histopathological effects in the kidney is considered *indeterminate*.

3.8.2. Synthesis of results for kidney effects—Taken together, the available evidence is inadequate to draw conclusions on the effects of DIBP on the kidney. Therefore, the evidence for kidney toxicity is considered *indeterminate* (Table S2).

3.9. Cancer

None of the available studies evaluated cancer in animals exposed to DIBP. The limited number of mutagenicity assays identified in the literature search generally had negative findings (Seed, 1982; Simmon et al., 1977; Zeiger et al., 1982), although genotoxicity assays of DIBP-treated primary human mucosal cells demonstrated DNA damage as measured by the comet assay (Kleinsasser et al., 2000a; Kleinsasser et al., 2000b; Kleinsasser et al., 2001). This was considered inadequate information to evaluate carcinogenicity; thus, the evidence for cancer is considered *indeterminate*.

4. Discussion

The results of this systematic review provide *robust* evidence that DIBP causes male reproductive and developmental toxicity and *slight* evidence for female reproductive toxicity and effects on liver, whereas evidence for effects on kidney and cancer were *indeterminate*. These results corroborate the CPSC (2014) and NAS (2017) conclusions that DIBP is a male reproductive toxic agent, with gestational exposure leading to permanent adverse effects in male offspring, and support DIBP as a children's health concern. The results also provide a reconnaissance of additional effects observed in laboratory animals after DIBP exposure, some of which may share the same mechanisms as the male reproductive effects.

Howdeshell et al. (2008) demonstrated that fetal testosterone production and fetal mortality (post-implantation loss) followed a similar dose-response relationship in rats exposed to a mixture of phthalates during gestation, and suggested that this might be because fetal T production and fetal mortality are both caused by decreased steroidogenesis (decreased testicular testosterone in male fetuses, and decreased ovarian progesterone in dams). While it has been found that DBP decreases maternal ovarian progesterone production in dams exposed at mid-pregnancy (Gray Jr et al., 2006), this has not yet been evaluated for DIBP. Gray Jr et al., 2006 hypothesized that decreased progesterone production could be mediated through decreased ovarian steroidogenesis, similar to the mechanism for decreased testicular steroidogenesis; or may result from disrupted androgen and INSL3 signaling, since those hormones contribute to maintaining the function of the corpora lutea in pregnant rats. There is also some epidemiologic evidence that DIBP is associated with decreased

expression of steroidogenic enzymes in the placenta (CYP11A1, CYP19A1, CYP17A1, 17 β -hydroxysteroid dehydrogenase) (Adibi et al., 2010) with a significantly stronger association in male placentas compared to female in some cases (Adibi et al., 2017), although results were inconsistent and suggested a possible non-monotonic response.

For male reproductive outcomes, the available studies suggest a sensitivity difference in mice compared to rats; however, there were relatively few mouse studies and all had some concerns for risk of bias and/or sensitivity, so it is difficult to fully delineate species sensitivity differences using the available data. Outcomes associated with the androgen-independent MOA for phthalates (e.g. atrophy of seminiferous cords, germ cell effects) were observed in both species. Conversely, while decreased T was observed in both species, mice did not consistently display the other androgen-dependent outcomes that were observed in rats (e.g. decreased testis weight and AGD). Similar observations on the relative sensitivity of mice and rats have been made for other phthalates, such as DBP. It has also been found that human xenografts (human fetal testis explants into rodents) are less sensitive than rats to the anti-androgenic effects of phthalates, suggesting an interspecies difference in sensitivity; however, the xenograft results have several factors that limit their interpretation, such as concerns about the developmental stage of the human tissues used in the experiments, and high variability in the results (Lioy et al., 2015). In the case of DIBP, results suggest that mice and rats are susceptible to both the androgen-dependent and -independent MOAs, although not all androgen-dependent biomarkers were conserved across species.

The available studies were generally not designed to evaluate female reproductive, liver, or kidney effects, so interpretation of these outcomes is limited. The few studies that evaluated female offspring following gestational exposure provide evidence of morphological effects (displaced ovaries and increased AGD), which could be caused by alterations in hormone signaling analogous to those seen in male offspring. While these observations are compelling, the implications for female fertility are unclear. Literature on the female reproductive effects of phthalates is relatively sparse. A review of phthalate effects in females reported altered steroidogenesis and ovarian function in animal studies, but effects were often inconsistent across studies and the mechanism was unclear (Kay et al., 2013). One study by Hannas et al. (2013) attempted to identify a critical window of exposure for in utero phthalate effects in the female rat, and found a statistically significant increase in vaginal and uterine agenesis in F1 females exposed to a phthalate mixture from GD 8–19; the incidence was lower after exposure from GD 8–13, and there were no female reproductive malformations after exposure from GD 14–19. One female exposed from GD 8–19 had undescended ovaries, located above the kidneys. Although the authors were not able to ascertain the critical window of exposure, they concluded that it appeared to encompass the major period of organogenesis (GD 8–13).

A concurrent systematic review of epidemiological studies of phthalates by our colleagues found *moderate* evidence of an association between DIBP exposure and reduced testosterone in adult cross-sectional studies, but only *slight* evidence of an association with other male reproductive effects. Comparatively, DBP and several other anti-androgenic phthalates had *robust* evidence of an association with male reproductive toxicity in humans (Radke et al., 2018). Given the effects of DIBP in animal models, the low level of evidence for

male reproductive toxicity in humans was likely due to the relatively small number and low sensitivity of the available epidemiological studies that evaluated DIBP exposure. Systematic review of the epidemiological literature also provided *moderate* evidence of an association between DIBP exposure and decreased birth size, and *slight* evidence for an association with preterm birth and spontaneous abortion (Radke et al., under review-a).

Epidemiological evidence also suggests that phthalate exposure may be associated with emerging health outcomes such as neurodevelopmental and metabolic toxicity, with *moderate* evidence of an association between DIBP exposure and diabetes risk, and *slight* evidence of an association with some neurodevelopmental outcomes (Radke et al., in preparation; Radke et al., under review-b). Although emerging health outcomes were not the focus of this systematic review of animal studies, we note that the literature search for DIBP identified one behavioral study (Ma et al., 2013), which reported that mice dosed with 1000 mg/kg-day DIBP via oral gavage for 8 weeks had decreased passive avoidance capability and increased apoptosis of hippocampal cells. We performed a preliminary evaluation of this study and found it to have reporting limitations that reduce confidence in the results (e.g., lack of experimental method detail to determine how memory and learning were measured in the passive avoidance test); however, it does provide suggestive evidence of neurological effects from DIBP exposure.

As a next step towards understanding the quantitative relationship between gestational DIBP exposure and phthalate syndrome in the male rat, an ordinal univariate dose-response analysis of phthalate syndrome was performed by our group using data from Saillenfait et al. (2008), in which endpoints were grouped by severity. This analysis will be available in a forthcoming publication.

This systematic review highlighted several ways in which future studies could provide further insight into the mechanisms and characterization of hazards of DIBP exposure. Most of the available studies were targeted at critical exposure windows for male reproductive toxicity, so it would be useful to design chronic or sub-chronic studies that cover critical windows of exposure for other outcomes. For instance, there are currently no one- or multi-generational reproductive toxicity studies available for DIBP; such a study would assess fertility and fecundity, and would provide a point of reference for comparing reproductive effects in animals exposed as adults versus those exposed in utero. Effects on liver, kidney, and cancer are known hazards for other phthalates, so further research is warranted to characterize these outcomes after exposure to DIBP, in addition to evaluating emerging outcomes such as neurodevelopmental toxicity. Additionally, despite the profound effects of DIBP on testosterone observed in males, data on hormone production in DIBP-exposed females was extremely limited. It would be useful to evaluate steroidogenesis and INSL3 in F1 females to provide a mechanistic understanding of how each sex is affected by gestational DIBP exposure, and to evaluate altered maternal steroidogenesis and INSL3 as a potential contributing factor to reduced fetal survival.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

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. We would like to acknowledge Anna Chen, Evangela Matthews, Swati Gummadi, Carolyn Gigot, Mefruz Haque (EPA student services contractors), and Carol Starkey (Oak Ridge Institute for Science and Education) for their assistance in data extraction and visualization for this systematic review. We would also like to thank Earl Gray, Justin Conley, and April Luke for their comments on previous drafts of this manuscript.

References

- Adibi JJ, Whyatt RM, Hauser R, Bhat HK, Davis BJ, Calafat AM, et al. , 2010. Transcriptional biomarkers of steroidogenesis and trophoblast differentiation in the placenta in relation to prenatal phthalate exposure. *Environ. Health Perspect* 118, 291–296. [PubMed: 20123604]
- Adibi JJ, Buckley JP, Lee MK, Williams PL, Just AC, Zhao Y, et al. , 2017. Maternal urinary phthalates and sex-specific placental mRNA levels in an urban birth cohort. *Environ. Health* 16, 35. [PubMed: 28381288]
- Albert O, Jégou B, 2014. A critical assessment of the endocrine susceptibility of the human testis to phthalates from fetal life to adulthood. *Hum. Reprod. Update* 20, 231–249. [PubMed: 24077978]
- 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]
- BASF, 2007. Diisobutylphthalate - Prenatal Developmental Toxicity Study in Wistar Rats Administration in the Diet (2007 update). Document Control Number: 86070000046. Submitted to the U.S. Environmental Protection Agency Under TSCA Section 8d.
- Boberg J, Metzдорff S, Wortzinger R, Axelstad M, Brokken L, Vinggaard AM, et al. , 2008. Impact of diisobutyl phthalate and other PPAR agonists on steroidogenesis and plasma insulin and leptin levels in fetal rats. *Toxicology* 250, 75–81. [PubMed: 18602967]
- Borch J, Axelstad M, Vinggaard AM, Dalggaard M, 2006. Diisobutyl phthalate has comparable anti-androgenic effects to di-*n*-butyl phthalate in fetal rat testis. *Toxicol. Lett* 163, 183–190. [PubMed: 16458459]
- Choy I, 1975. Effects of sulfathiazole, phthalylsulfathiazole and phthalic acid esters on the development of chick embryos [translated]. *Showa Igakkai Zasshi* 35, 187–201.
- Consumer Product Safety Commission (CPSC), 2014. Chronic Hazard Advisory Panel on Phthalates and Phthalate Alternatives (with Appendices). U.S. Consumer Product Safety Commission, Directorate for Health Sciences, Bethesda, MD.
- 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.
- Eastman Kodak, 1954. Screening test on the toxicity of diisobutyl phthalate. In: EPA/OTS Doc #878214403. Rochester, NY.
- Elsisi AE, Carter DE, Sipes IG, 1989. Dermal absorption of phthalate diesters in rats. *Fundam. Appl. Toxicol* 12, 70–77. [PubMed: 2925020]
- Foster P, Gray LE Jr. (Eds.), 2013. *Toxic Responses of the Reproductive System*. McGraw-Hill Education, New York, NY.
- Foster PM, Lake BG, Thomas LV, Cook MW, Gangolli SD, 1981. Testicular effects and zinc excretion produced by various isomers of monobutyl-*o*-phthalate in the rat. *Chem. Biol. Interact* 34, 233–238. [PubMed: 7460085]
- Foster PM, Lake BG, Cook MW, Thomas LV, Gangolli SD (Eds.), 1982. *Structure-Activity Requirements for the Induction of Testicular Atrophy by Butyl Phthalates in Immature Rats: Effect on Testicular Zinc Content*. Plenum Press, New York, NY.
- Fromme H, Gruber L, Seckin E, Raab U, Zimmermann S, Kiranoglu M, et al. , 2011. Phthalates and their metabolites in breast milk—results from the Bavarian Monitoring of Breast Milk (BAMBI). *Environ. Int* 37, 715–722. [PubMed: 21406311]

- 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., Laskey J, Ostby J, 2006. Chronic di-*n*-butyl phthalate exposure in rats reduces fertility and alters ovarian function during pregnancy in female Long-Evans hooded rats. *Toxicol. Sci* 93, 189–195. [PubMed: 16763070]
- 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]
- Hannas BR, Lambright CS, Furr J, Howdeshell KL, Wilson VS, Gray LE, 2011. Dose-response assessment of fetal testosterone production and gene expression levels in rat testes following in utero exposure to diethylhexyl phthalate, diisobutyl phthalate, diisooheptyl phthalate, and diisononyl phthalate. *Toxicol. Sci* 123, 206–216. [PubMed: 21633115]
- Hannas BR, Lambright CS, Furr J, Evans N, Foster PM, Gray EL, et al. , 2012. Genomic biomarkers of phthalate-induced male reproductive developmental toxicity: a targeted rt-PCR array approach for defining relative potency. *Toxicol. Sci* 125, 544–557. [PubMed: 22112501]
- Hannas BR, Howdeshell KL, Furr J, Earl Gray L Jr., 2013. In utero phthalate effects in the female rat: a model for MRKH syndrome. *Toxicol. Lett* 223, 315–321. [PubMed: 23542816]
- 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, 1983a. Final Report. Screening of Priority Chemicals for Potential Reproductive Hazard NIOSH/00133459. OH:NIOSH (The National Institute for Occupational Safety and Health), Cincinnati.
- Hazleton Laboratories, 1983b. Screening of Priority Chemicals for Potential Reproductive Hazard (Final Report) With Attachments and Cover Sheet. TSCATS/400290, OTS0516205, Doc I.D. 86–870001624. Shell Oil Company, Houston, TX.
- Hazleton Laboratories, 1992. Initial Submission: Screening of Priority Chemicals for Potential Reproductive Hazard (Final Report) With Cover Letter Dated 041492. 88920002015. U.S. Environmental Protection Agency, Washington, DC.
- Higgins JPT, 2011. Chapter 8: Assessing Risk of Bias in Included Studies. In: *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0. 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]
- Hotchkiss AK, Rider CV, Blystone CR, Wilson VS, Hartig PC, Ankley GT, et al. , 2008. Fifteen years after “Wingspread”—environmental endocrine disruptors and human and wildlife health: where we are today and where we need to go. *Toxicol. Sci* 105, 235–259. [PubMed: 18281716]
- 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]
- HSDB, 2017. Diisobutyl Phthalate. National Library of Medicine, Bethesda, MD.
- 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]
- Kay VR, Chambers C, Foster WG, 2013. Reproductive and developmental effects of phthalate diesters in females. *Crit. Rev. Toxicol* 43, 200–219. [PubMed: 23405971]
- Kleinsasser NH, Kastenbauer ER, Weissacher H, Muenzenrieder RK, Harreus UA, 2000a. Phthalates demonstrate genotoxicity on human mucosa of the upper aero-digestive tract. *Environ. Mol. Mutagen* 35, 9–12.

- Kleinsasser NH, Weissacher H, Kastenbauer ER, Dirschedl P, Wallner BC, Harreus UA, 2000b. Altered genotoxicity in mucosal cells of head and neck cancer patients due to environmental pollutants. *Eur. Arch. Otorhinolaryngol* 257, 337–342. [PubMed: 10993555]
- Kleinsasser NH, Wallner BC, Kastenbauer ER, Weissacher H, Harreus UA, 2001. Genotoxicity of di-butyl-phthalate and di-iso-butyl-phthalate in human lymphocytes and mucosal cells. *Teratog. Carcinog. Mutagen* 21, 189–196. [PubMed: 11301413]
- Koch HM, Christensen KLY, Harth V, Lorber M, Brüning T, 2012. Di-*n*-butyl phthalate (DnBP) and diisobutyl phthalate (DIBP) metabolism in a human volunteer after single oral doses. *Arch. Toxicol* 86, 1829–1839. [PubMed: 22820759]
- Latini G, Wittassek M, Del Vecchio A, Presta G, De Felice C, Angerer J, 2009. Lactational exposure to phthalates in southern Italy. *Environ. Int* 35, 236–239. [PubMed: 18684505]
- Lawrence WH, Malik M, Turner JE, Singh AR, Autian J, 1975. A toxicological investigation of some acute, short-term, and chronic effects of administering di-2-e-thylhexyl phthalate (DEHP) and other phthalate esters. *Environ. Res* 9, 1–11. [PubMed: 1122902]
- Lioy PJ, Hauser R, Gennings C, Koch HM, Mirkes PE, Schwetz BA, et al. , 2015. Assessment of phthalates/phthalate alternatives in children's toys and childcare articles: review of the report including conclusions and recommendation of the chronic Hazard advisory panel of the consumer product safety commission. *J. Expo. Sci. Environ. Epidemiol* 25, 343–353. [PubMed: 25944701]
- Ma N, Liu S, Gao P, Cao P, Xu H, 2013. Effect of diisobutyl phthalate on learning and memory behavior and apoptosis of hippocampus cells in mice [translated]. *Wei Sheng Yan Jiu* 42, 57–60. [PubMed: 23596708]
- Martino-Andrade AJ, Chahoud I, 2010. Reproductive toxicity of phthalate esters. *Mol. Nutr. Food Res* 54, 148–157. [PubMed: 19760678]
- National Academy of Sciences (NAS), 2017. Application of Systematic Review Methods in an Overall Strategy for Evaluating Low-Dose Toxicity from Endocrine Active Chemicals. The National Academies Press, Washington, D.C.
- National Research Council, 2008. Phthalates and Cumulative Risk Assessment: The Task Ahead. National Academies Press, Washington, DC.
- Noriega NC, Howdeshell KL, Furr J, Lambright CR, Wilson VS, Gray LE Jr., 2009. Pubertal administration of DEHP delays puberty, suppresses testosterone production, and inhibits reproductive tract development in male Sprague-Dawley and Long-Evans rats. *Toxicol. Sci* 111, 163–178. [PubMed: 19528224]
- Oishi S, Hiraga K, 1980a. Effect of phthalic acid esters on mouse testes. *Toxicol. Lett* 5, 413–416. [PubMed: 7394838]
- Oishi S, Hiraga K, 1980b. Testicular atrophy induced by phthalic acid esters: effect on testosterone and zinc concentrations. *Toxicol. Appl. Pharmacol* 53, 35–41. [PubMed: 7385236]
- Oishi S, Hiraga K, 1980c. Testicular atrophy induced by phthalic acid monoesters: effects of zinc and testosterone concentrations. *Toxicology* 15, 197–202. [PubMed: 7466832]
- Oishi S, Hiraga K, 1980d. Effects of phthalic acid monoesters on mouse testes. *Toxicol. Lett* 6, 239–242. [PubMed: 7423552]
- Radke E, Nachman R, Braun J, Cooper G, Phthalate exposure and neuropsychological and behavioral effects: a systematic review and meta-analysis of the human epidemiological evidence (in preparation).
- Radke E, Glenn B, Braun J, Cooper G, Phthalate exposure and female reproductive and developmental effects: a systematic review of the human epidemiological evidence (under review-a).
- Radke E, Galizia A, Thayer K, Cooper G Phthalate exposure and metabolic effects: a systematic review of the human epidemiological evidence (under review-b).
- 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 (Pt 1), 764–793. 10.1016/j.envint.2018.07.029. Epub 2018 Oct 16. [PubMed: 30336412]
- Ray B, D' Souza AS, Kumar V, Pugazhandhi B, D'Souza MR, Nayak D, et al. , 2012. Ovarian development in Wistar rat treated prenatally with single dose diisobutyl phthalate. *Bratisl. Lek. Listy* 113, 577–582. [PubMed: 23094893]

- Saillenfait AM, Sabate JP, Gallissot F, 2006. Developmental toxic effects of diisobutyl phthalate, the methyl-branched analogue of di-*n*-butyl phthalate, administered by gavage to rats. *Toxicol. Lett* 165, 39–46. [PubMed: 16516415]
- Saillenfait AM, Sabaté JP, Gallissot F, 2008. Diisobutyl phthalate impairs the androgen-dependent reproductive development of the male rat. *Reprod. Toxicol* 26, 107–115. [PubMed: 18706996]
- Saillenfait AM, Sabaté JP, Denis F, Antoine G, Robert A, Roudot AC, et al. , 2017. Evaluation of the effects of α -cypermethrin on fetal rat testicular steroidogenesis. *Reprod. Toxicol*
- Schunemann H, Hill S, Guyatt G, Akl EA, Ahmed F, 2011. The GRADE approach and Bradford Hill's criteria for causation. *J. Epidemiol. Community Health* 65, 392–395. [PubMed: 20947872]
- Sedha S, Gautam AK, Verma Y, Ahmad R, Kumar S, 2015. Determination of in vivo estrogenic potential of di-isobutyl phthalate (DIBP) and di-isononyl phthalate (DINP) in rats. *Environ. Sci. Pollut. Res. Int* 22, 18197–18202. [PubMed: 26178834]
- Seed JL, 1982. Mutagenic activity of phthalate esters in bacterial liquid suspension assays. *Environ. Health Perspect* 45, 111–114. [PubMed: 6754360]
- Simmon VF, Kauhanen K, Tardiff RG (Eds.), 1977. *Mutagenic Activity of Chemicals Identified in Drinking Water*. Elsevier/North Holland Press, New York, NY.
- Singh AR, Lawrence WH, Autian J, 1972. Teratogenicity of phthalate esters in rats. *J. Pharm. Sci* 61, 51–55. [PubMed: 5058645]
- Sohn J, Kim S, Koschorreck J, Kho Y, Choi K, 2016. Alteration of sex hormone levels and steroidogenic pathway by several low molecular weight phthalates and their metabolites in male zebrafish (*Danio rerio*) and/or human adrenal cell (H295R) line. *J. Hazard. Mater* 320, 45–54. [PubMed: 27513369]
- Strucinski P, Goralczyk K, Ludwicki JK, Czaja K, Hernik A, Korcz W, 2006. Levels of selected organochlorine insecticides, polychlorinated biphenyls, phthalates and perfluorinated aliphatic substances in blood. *Rocz. Panstw. Zakl. Hig* 57, 99–112. [PubMed: 17044303]
- University of Rochester, 1953. One month feeding tests of di-isobutyl phthalate with cover letter. In: EPA/OTS Doc #878212229. University of Rochester School of Medicine and Dentistry, Rochester, NY.
- University of Rochester, 1954. Preliminary acute toxicity tests and short term feeding tests of rats and dogs given di-isobutyl phthalate and di-butyl phthalate. In: EPA/OTS Doc #878210833. University of Rochester School of Medicine and Dentistry, Rochester, NY.
- US EPA, 1988. Recommendations for and documentation of biological values for use in risk assessment. In: EPA/600/6-87/008. U.S. Environmental Protection Agency, Office of Research and Development, Office of Health and Environmental Assessment, Cincinnati, OH.
- US EPA, 1991. Guidelines for developmental toxicity risk assessment. In: EPA/600/FR-91/001. U.S. Environmental Protection Agency, Risk Assessment Forum, Washington, DC.
- US EPA, 2006. A framework for assessing health risk of environmental exposures to children. In: EPA/600/R-05/093F. U.S. Environmental Protection Agency, Office of Research and Development, National Center for Environmental Assessment, Washington, DC.
- US EPA, 2007. Integrated Summary Report for Validation of a Test Method for Assessment of Pubertal Development and Thyroid Function in Juvenile Male Rats as a Potential Screen in the Endocrine Disruptor Screening Program Tier-1 Battery. U.S. Environmental Protection Agency Office of Science Coordination and Policy and Office of Research and Development, Washington, D.C.
- Wang X, Sheng N, Cui R, Zhang H, Wang J, Dai J, 2017. Gestational and lactational exposure to di-isobutyl phthalate via diet in maternal mice decreases testosterone levels in male offspring. *Chemosphere* 172, 260–267. [PubMed: 28081510]
- Wittassek M, Wiesmuller GA, Koch HM, Eckard R, Dobler L, Muller J, et al. , 2007. Internal phthalate exposure over the last two decades—a retrospective human biomonitoring study. *Int. J. Hyg. Environ. Health* 210, 319–333. [PubMed: 17400024]
- Wittassek M, Angerer J, Kolossa-Gehring M, Schafer SD, Klockenbusch W, Dobler L, et al. , 2009. Fetal exposure to phthalates—a pilot study. *Int. J. Hyg. Environ. Health* 212, 492–498. [PubMed: 19423389]

- Wormuth M, Scheringer M, Vollenweider M, Hungerbühler K, 2006. What are the sources of exposure to eight frequently used phthalic acid esters in Europeans? *Risk Anal.* 26, 803–824. [PubMed: 16834635]
- Zeiger E, Haworth S, Speck W, Mortelmans K, 1982. Phthalate ester testing in the National Toxicology Program's environmental mutagenesis test development program. *Environ. Health Perspect* 45, 99–101. [PubMed: 6754365]
- Zhu XB, Tay TW, Andriana BB, Alam MS, Choi EK, Tsunekawa N, et al. , 2010. Effects of di-isobutyl phthalate on testes of prepubertal rats and mice. *Okajimas Folia Anat. Jpn* 86, 129–136. [PubMed: 20560449]
- 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]

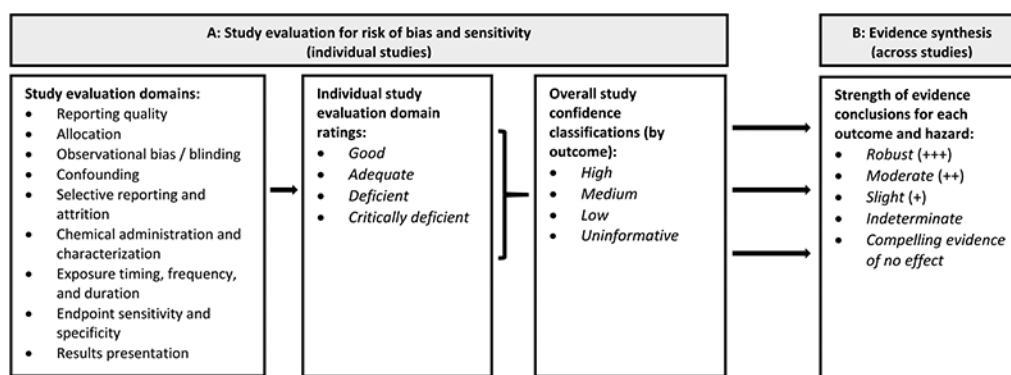
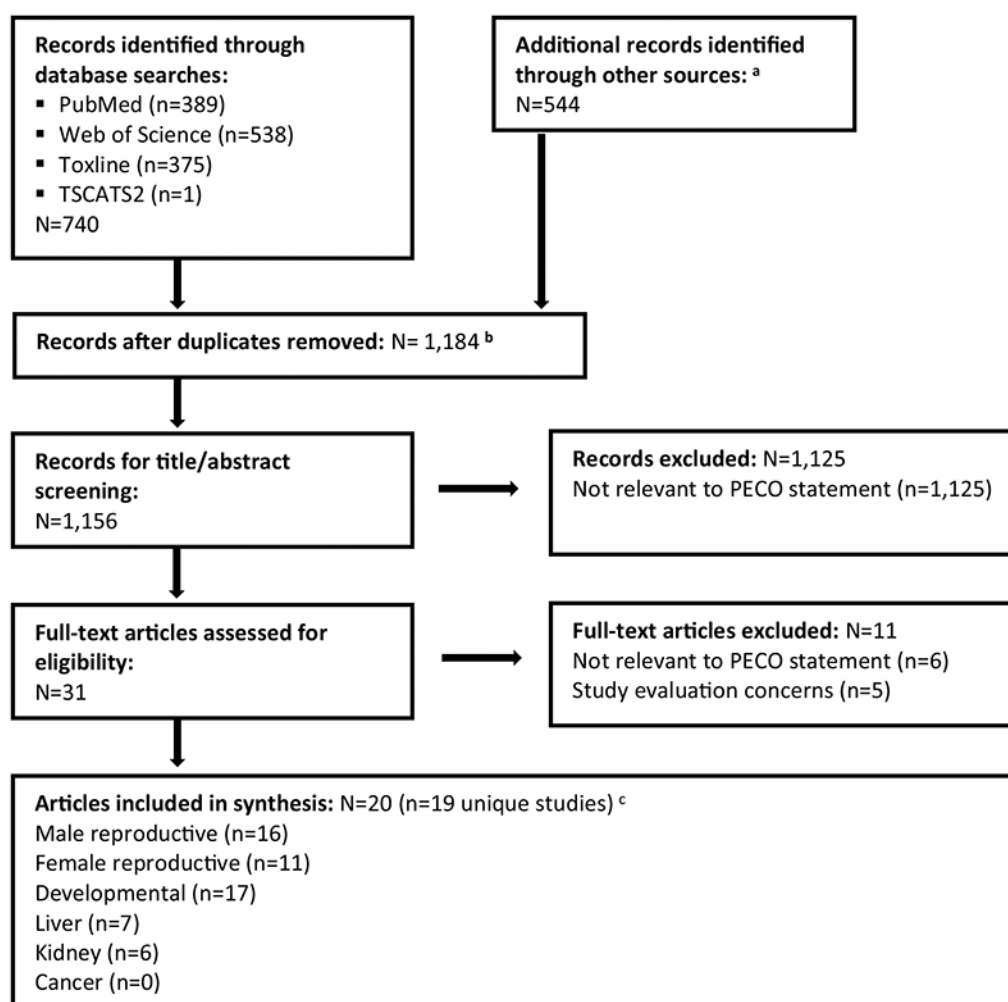


Fig. 1. Summary of (A) study evaluation and (B) strength of evidence characterization for DIBP animal toxicology studies.

**Fig. 2.**

Literature flow diagram for identifying DIBP animal toxicology studies.

^a Other sources consisted of forward and backward searches, searching citations from key references, manual search of citations from key regulatory documents, references that had been previously identified from an earlier DIBP review effort, and supplementary materials for articles identified during the literature search.

^b Includes 28 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.

^c Most studies reported data on multiple hazards; see Table 1.

Table 1

Summary of studies, overall study confidence classifications by outcome^a.

Author (year)	Species (strain)	Exposure life stage and duration	Exposure route	Male reproductive ^b			Female reproductive ^c		
				Testosterone	Morphological development	Reproductive organ weight	Testicular histology and sperm	Morphological development	Maternal body weight
BASF, 2007	Rat (Wistar)	GD 6-20	Diet	-	-	-	-	-	H
Borch et al., 2006	Rat (Wistar)	GD 7-19; GD 7-20/21	Gavage	H	H	-	M	H	H
Saillenfait et al., 2006	Rat (Sprague-Dawley)	GD 6-20	Gavage	-	H	-	-	H	H
Howdeshell et al., 2008	Rat (Sprague-Dawley)	GD 8-18	Gavage	H	-	-	-	-	M
Saillenfait et al., 2008	Rat (Sprague-Dawley)	GD 12-21	Gavage	-	H	H	H	H	M
Saillenfait et al., 2017	Rat (Sprague-Dawley)	GD 13-19	Gavage	H	H	-	-	-	H
Furr et al., 2014	Rat (Sprague-Dawley)	GD 14-18	Gavage	H	-	-	-	-	L
Hannas et al., 2012	Rat (Sprague-Dawley)	GD 14-18	Gavage	H	-	-	-	-	L
Hannas et al., 2011	Rat (Sprague-Dawley)	GD 14-18	Gavage	H	-	-	-	-	L
Wang et al., 2017	Mouse (ICR)	GD 0-21; GD 0-PND 21	Diet	M	M	M	M	-	L
Sedha et al., 2015	Rat (Wistar)	PND 21-23; PND 21-40	Gavage	-	-	-	-	M	-
Zhu et al., 2010	Rat (Sprague-Dawley); Mouse (C57B1/6N)	PND 21-28	Gavage	-	-	L	-	-	-
Oishi and Hiraga, 1980a	Mouse (JCL:ICR)	~PND 35-42	Diet	M	-	L	-	-	-
Oishi and Hiraga, 1980b	Rat (JCL:Wistar)	~PND 35-42	Diet	L	-	M	L	-	-
Oishi and Hiraga, 1980c [*]	Rat (JCL:Wistar)	~PND 35-42	Diet	M	-	M	-	-	-
Oishi and Hiraga, 1980d [*]	Mouse (JCL:ICR)	~PND 35-42	Diet	M	-	L	-	-	-
Foster et al., 1981, 1982 [*]	Rat (Sprague-Dawley)	"Young"; 6-day exposure	Gavage	-	-	L	L	-	-

Author (year)	Species (strain)	Exposure life stage and duration	Exposure route	Male reproductive ^b				Female reproductive ^c		
				Testosterone	Morphological development	Reproductive organ weight	Testicular histology and sperm	Morphological development	Maternal body weight	
University of Rochester, 1953	Rat (albino; strain not reported)	Weaning to 1 month post-weaning	Diet	-	-	-	-	-	-	-
University of Rochester, 1954	Rat (albino; strain not reported)	Weaning to 4 months post-weaning	Diet	-	-	M	-	-	-	-
Female reproductive^c				Developmental^d				Kidney		
Author (year)	Gestation length	Reproductive organ weight	Survival	Growth	Structural alterations	Organ weight	Histopathology	Organ weight	Histopathology	
BASF, 2007	-	H	H	H	H	-	-	-	-	-
Borch et al., 2006	-	-	H	H	-	-	-	-	-	-
Saillenfait et al., 2006	-	H	H	H	H	-	-	-	-	-
Howdeshell et al., 2008	-	-	H	-	-	-	-	-	-	-
Saillenfait et al., 2008	H	-	H	H	-	-	-	-	-	-
Saillenfait et al., 2017	-	H	H	H	H	-	-	-	-	-
Furr et al., 2014	-	-	M	-	-	-	-	-	-	-
Hannas et al., 2012	-	-	M	-	-	-	-	-	-	-
Hannas et al., 2011	-	-	L	-	-	-	-	-	-	-
Wang et al., 2017	-	-	H	M	-	M	-	-	-	-
Sedha et al., 2015	-	M	-	M	-	-	-	-	-	-
Zhu et al., 2010	-	-	-	-	-	-	-	-	-	-
Oishi and Hiraga, 1980a	-	-	-	L	-	M	-	M	-	-
Oishi and Hiraga, 1980b	-	-	-	L	-	M	-	M	-	-
Oishi and Hiraga, 1980c [*]	-	-	-	L	-	-	-	-	-	-
Oishi and Hiraga, 1980d [*]	-	-	-	L	-	M	-	M	-	-
Foster et al., 1981, 1982 [*]	-	-	-	-	-	L	-	L	-	-
University of Rochester, 1953	-	-	-	L	-	M	L	M	L	L
University of Rochester, 1954	-	-	-	L	-	M	L	M	L	L

^{*} indicates MIBP study

GD = Gestation day, PND = Postnatal day

^aHigh confidence (H), Medium confidence (M), Low confidence (L). Dash (-) indicates outcomes that were not included in a study.

^bMale reproductive outcomes: Testosterone (testicular production or level measured in testis or serum), morphological development (AGD, nipple retention, hypospadias, cryptorchidism, cleft prepuce, time to puberty), testicular histology and sperm (sperm counts, motility, morphology, histological evaluations of testicular atrophy or azoospermia/oligospermia), reproductive organ weights (testis, epididymides, prostate, seminal vesicles)

^cFemale reproductive outcomes: Maternal body weight (body weight gain during gestation or lactation), gestation length, morphological development (AGD, displaced ovaries, time to puberty), reproductive organ weight (uterus, vagina, ovary)

^dDevelopmental outcomes: Survival (fetal viability, fetal mortality, resorptions, pre- or post-implantation loss), growth (pre- or postnatal body weight), malformations/variations (external, skeletal, soft tissue/visceral)

Table 2

Evidence profile table for male reproductive effects of DIBP or MIBP

Male Reproductive Effects				
Outcome	Available studies	Factors that increase confidence	Factors that decrease confidence	Confidence judgement for outcome
Gestational (F1) Exposure				
Testosterone	High confidence: Borch et al., 2006 Furr et al., 2014 Hannas et al., 2011 Hannas et al., 2012 Howdeshell et al., 2008 Saillenfait et al., 2017 Medium confidence: Wang et al., 2017	<ul style="list-style-type: none"> Consistency Dose-response gradient Effect size Biological plausibility (support from mechanistic evidence) Minimal concerns for bias and sensitivity 		⊕⊕⊕ ROBUST Supported by consistency and coherence across outcomes, with mechanistic evidence (e.g. decreased testicular expression of steroidogenic enzymes and INSL3 in F1 males) providing support for biological plausibility. The greatest weight of evidence came from gestational exposure studies, whereas postnatal exposure studies were limited by risk of bias and Male sensitivity concerns.
Male morphological development	High confidence: Borch et al., 2006 Saillenfait et al., 2006 Saillenfait et al., 2008 Saillenfait et al., 2017 Medium confidence: Wang et al., 2017	<ul style="list-style-type: none"> Consistency within rat studies Dose-response gradient Effect size Biological plausibility Minimal concerns for bias and sensitivity 		⊕⊕⊕ ROBUST All rat studies observed a dose-related increase in effects consistent with decreased T and INSL3, including increased time to puberty, decreased AGD, nipple retention, cryptorchidism, nonscrotal testis, hypospadias, and exposed os penis. No effects on AGD were observed in mice (Wang et al., 2017).
Reproductive organ weight	High confidence: Saillenfait et al., 2008 Medium confidence: Wang et al., 2017	<ul style="list-style-type: none"> Dose-response gradient and minimal concern for bias and sensitivity in the study by Saillenfait et al., 2008 Biological plausibility 	<ul style="list-style-type: none"> Few studies 	⊕⊕○ MODERATE Decreased reproductive organ weights were observed in rats (Saillenfait et al., 2008), whereas a consistent trend in testis weight was not observed in mice (Wang et al., 2017); it is not clear whether this is related to differences in species or study designs. This effect is consistent with decreased T.

Male Reproductive Effects				
Outcome	Available studies	Factors that increase confidence	Factors that decrease confidence	Confidence judgement for outcome
		• Coherence with postnatal exposure studies		Confidence judgement for overall hazard
Testicular histology or sperm evaluation	High confidence: Saillenfiat et al., 2008	• Consistency		⊕⊕⊕ ROBUST Adverse effects on the testis and/or sperm were observed in rats and mice, including a dose-related increased incidence of pathological lesions of the testis (Borch et al., 2006; Saillenfiat et al., 2008), epididymal oligo- or azoospermia (Saillenfiat et al., 2008), and decreased sperm concentration and motility (Wang et al., 2017).
	Medium confidence: Borch et al., 2006 Wang et al., 2017	• Dose-response gradient • Effect size • Biological plausibility		
Postnatal (Weanling or Peripubertal) Exposure	Medium confidence: Oishi and Hiraga, 1980a Oishi and Hiraga, 1980c * Oishi and Hiraga, 1980d *	• Biological plausibility	• Concerns for bias and sensitivity in some studies • Unexplained inconsistency	○○○ INDETERMINATE A dose-related increase in androgen levels was observed in two rat studies (Oishi and Hiraga, 1980b, c), whereas androgen levels were decreased or not changed in two mouse studies (Oishi and Hiraga, 1980a, d).
	Low confidence: Oishi and Hiraga, 1980b			
Reproductive organ weight	Medium confidence: Oishi and Hiraga, 1980b Oishi and Hiraga, 1980c * U. Rochester, 1954	• Consistency within rat studies • Biological plausibility	• Concerns for bias and sensitivity in some studies	⊕⊕⊕ ROBUST In rats, a dose-related decrease in absolute testis weight was consistently observed (Oishi and Hiraga, 1980b, c; Foster et al., 1981; University of Rochester, 1954). In weanling mice, Zhu et al. (2010) observed decreased absolute testis weight in the highest dose group. In peripubertal mice, Oishi and Hiraga (1980a, d) observed increased relative testis weight, which is considered a less reliable metric compared to absolute testis weight.
	Low confidence: Oishi and Hiraga, 1980a Oishi and Hiraga, 1980d * Foster et al., 1981 Zhu et al., 2010	• Coherence with gestational exposure studies		

Male Reproductive Effects					
Outcome	Available studies	Factors that increase confidence	Factors that decrease confidence	Confidence judgement for outcome	Confidence judgement for overall hazard
Testicular histology/sperm evaluation	Low confidence: Oishi and Hiraga, 1980b Foster et al., 1981*	<ul style="list-style-type: none"> Consistency Biological plausibility Coherence with gestational exposure studies 	<ul style="list-style-type: none"> Concerns for bias and sensitivity Few studies 	⊕⊕○ MODERATE Rats were found to have increased testicular atrophy (Foster et al., 1981) and decreased spermatocytes and spermatogonia (Oishi and Hiraga, 1980a).	

* indicates MIBP study

Table 3

Evidence profile table for female reproductive effects of DIBP

Female Reproductive Effects					
Outcome	Available studies	Factors that increase confidence	Factors that decrease confidence	Confidence judgement for outcome	Confidence judgement for overall hazard
<i>Gestational (F1) or Postnatal (Weanling) Exposure</i>	Female morphological development				
	High confidence: Saillenfait et al., 2006 Saillenfait et al., 2008 Borch et al., 2006	• Biological plausibility (coherence with effects observed in males)	• Concerns for study sensitivity	⊕○○○ SLIGHT	⊕○○○ SLIGHT
	Medium confidence: Sedha et al., 2015			Increased female AGD was observed, but was not always statistically significant (Borch et al., 2006, Saillenfait et al., 2008). A non-significant increase in displaced ovaries was observed by Saillenfait et al., 2006. DIBP did not accelerate female puberty in a 20-day pubertal assay (Sedha et al., 2015).	Based on limited evidence for effects on female AGD and ovary displacement in F1 females following gestational exposure, and for effects on maternal weight gain in F0 females following maternal exposure. There were concerns for study sensitivity because most gestational exposure studies were designed to evaluate male reproductive effects, and the exposure windows may not have been the most sensitive for detecting F1 female or maternal toxicity.
Maternal (F0) Exposure	Reproductive organ weight				
	Medium confidence: Sedha et al., 2015		• Single study	○○○○ INDETERMINATE	No effects on were observed on uterus, ovary, or vagina weight in 3- or 20-day assays in prepubescent females.
	Maternal body weight	• Dose-response gradient and minimal concern for bias in the study by BASF, 2007	• Concerns for study sensitivity	⊕○○○ SLIGHT	Corrected maternal body weight was significantly decreased in one study (BASF, 2007). Otherwise, any effects on maternal weight gain were concurrent with decreased gravid uterine weight or decreased offspring body weight, and therefore appeared to be secondary effects related to fetal toxicity.
Reproductive organ weight	High confidence: BASF, 2007 Saillenfait et al., 2006 Saillenfait et al., 2017				
	Medium confidence: Saillenfait et al., 2008 Howdeshell et al., 2008				
	Low confidence: Furr et al., 2014 Hannas et al., 2011 Hannas et al., 2012 Wang et al., 2017				
Reproductive organ weight	High confidence: BASF, 2007 Saillenfait et al., 2006 Saillenfait et al., 2006	• Minimal concerns for bias	• Concerns for study sensitivity	○○○○ INDETERMINATE	Saillenfait et al., 2006 reported decreased gravid uterine weight, which appeared to be a secondary effect related to fetal toxicity. Otherwise, no effects on female

Female Reproductive Effects					
Outcome	Available studies	Factors that increase confidence	Factors that decrease confidence	Confidence judgement for outcome	Confidence judgement for overall hazard
Gestation length	High confidence: Saillenfaït et al., 2008	<ul style="list-style-type: none">Minimal concern for bias	<ul style="list-style-type: none">Single studyConcerns for study sensitivity	reproductive organ weight were observed. ○○○ INDETERMINATE Effects on gestation length were not observed in one study that began exposing dams at mid-gestation, which may not be the most sensitive window of exposure for this endpoint	

Table 4

Evidence profile table for developmental effects of DIBP or MIBP

Developmental effects				
Outcome	Available studies	Factors that increase confidence	Factors that decrease confidence	Confidence judgement for overall hazard
Gestational (F1) Exposure				
Fetal survival	<p>High confidence: BASF, 2007; Borch et al., 2006; Howdeshell et al., 2008; Saillenfait et al., 2006; Saillenfait et al., 2008; Saillenfait et al., 2017; Wang et al., 2017</p> <p>Medium confidence: Hannas et al., 2012; Furr et al., 2014</p> <p>Low confidence: Hannas et al., 2011</p>	<ul style="list-style-type: none"> Dose-response gradient, effect size, and minimal concern for bias in studies by Saillenfait et al., 2006 and Howdeshell et al., 2008 Biological plausibility (support from mechanistic evidence for DBP) 	<p>• Few studies</p>	<p>⊕⊕⊕ ROBUST</p> <p>A dose-related increase in post-implantation loss was observed in two Sprague-Dawley rat studies that exposed animals from GD 6-20 and GD 8-18, respectively (Howdeshell et al., 2008; Saillenfait et al., 2006), but not in Wistar rats or mice exposed for a similar duration. Fetal survival was also not affected in studies that exposed Sprague-Dawley rats for shorter durations in mid- to late gestation.</p> <p>⊕⊕⊕ ROBUST</p> <p>Based on consistent evidence of reduced fetal and postnatal growth across studies, evidence of reduced fetal survival in Sprague-Dawley rats exposed for longer durations in gestation, and slight evidence of structural alterations following gestational exposure.</p>
Fetal growth	<p>High confidence: BASF, 2007; Borch et al., 2006; Saillenfait et al., 2006; Saillenfait et al., 2008; Saillenfait et al., 2017; Wang et al., 2017</p> <p>Medium confidence: Hannas et al., 2011</p>	<ul style="list-style-type: none"> Consistency Dose-response gradient Effect size Minimal concern for bias 	<p>•</p>	<p>⊕⊕⊕ ROBUST</p> <p>A dose-related decrease in fetal body weights or body weights at birth were consistently in rats at higher dose levels, but were not observed in lower dose studies.</p>
Fetal structural alterations	<p>High confidence: BASF, 2007; Saillenfait et al., 2006; Saillenfait et al., 2017</p>	<ul style="list-style-type: none"> Dose-response gradient and minimal concern for bias and sensitivity in the study by Saillenfait et al., 2006 	<p>• Few studies</p>	<p>⊕○○ SLIGHT</p> <p>A dose-related increase in external, visceral, and skeletal malformations was reported in Sprague-Dawley rats by Saillenfait et al., 2006, but not in a similar study in Wistar rats by BASF, 2007. There were no external malformations in a lower dose study by Saillenfait et al., 2017.</p>
Postnatal survival	<p>High confidence: Saillenfait et al., 2008</p>	<ul style="list-style-type: none"> Minimal concern for bias 	<p>• Single study</p>	<p>○○○ INDETERMINATE</p> <p>Effects on postnatal survival following gestational exposure were not observed in the single study that evaluated this outcome.</p>
Postnatal growth	<p>High confidence:</p>	<ul style="list-style-type: none"> Dose-response gradient and 	<p>• Few studies</p>	<p>⊕⊕○ MODERATE</p>

Developmental effects				
Outcome	Available studies	Factors that increase confidence	Factors that decrease confidence	Confidence judgement for outcome Confidence judgement for overall hazard
Postnatal (Weanling or Peripubertal) Exposure	Saillenfait et al., 2008	minimal concern for bias in the study by Saillenfait et al., 2008		A dose-related decrease in postnatal growth was observed in rats exposed during gestation (Saillenfait et al., 2008), but not in a mouse study that tested a single, lower dose of DIBP (Wang et al., 2017). ⊕⊕○ MODERATE A dose-related decrease in postnatal and adult growth was observed in all available peripubertal exposure studies in rats and mice. However, in the low confidence studies, the reduced growth may have been a secondary effect related to reduced food consumption.
	Medium confidence: Wang et al., 2017			
	Postnatal growth	• Consistency	• Concerns for bias and sensitivity in some studies	
	Medium confidence: Sedha et al., 2015	• Dose-response gradient		
	Low confidence: Oishi and Hiraga, 1980a Oishi and Hiraga, 1980b Oishi and Hiraga, 1980c * Oishi and Hiraga, 1980d * U. Rochester, 1953 U. Rochester, 1954	• Effect size		

* indicates MIBP study