



EPA Public Access

Author manuscript

Environ Int. Author manuscript; available in PMC 2022 August 24.

About author manuscripts

Submit a manuscript

Published in final edited form as:

Environ Int. 2019 September ; 130: 104580. doi:10.1016/j.envint.2019.02.003.

Phthalate exposure and female reproductive and developmental outcomes: a systematic review of the human epidemiological evidence

Elizabeth G. Radke^{a,*}, Barbara S. Glenn^a, Joseph M. Braun^b, Glinda S. Cooper^{a,c}

^aU.S. Environmental Protection Agency, National Center for Environmental Assessment, United States

^bBrown University, School of Public Health, United States

^cThe Innocence Project, United States

Abstract

Objective: We performed a systematic review of the epidemiology literature to identify the female reproductive and developmental effects associated with phthalate exposure.

Data sources and study eligibility criteria: Six phthalates were included in the review: di(2-ethylhexyl) phthalate (DEHP), diisononyl phthalate (DINP), dibutyl phthalate (DBP), diisobutyl phthalate (DIBP), butyl benzyl phthalate (BBP), and diethyl phthalate (DEP). The initial literature search (of PubMed, Web of Science, and Toxline) included all studies of female reproductive and developmental effects in humans, and outcomes were selected for full systematic review based on data availability.

Study evaluation and synthesis methods: For each outcome, studies were evaluated using criteria defined a priori for risk of bias and sensitivity by two reviewers using a domain-based approach. Evidence was synthesized by outcome and phthalate and strength of evidence was summarized using a structured framework.

Results: The primary outcomes reviewed here are (number of included/excluded studies in parentheses): pubertal development (5/13), time to pregnancy (3/4), preterm birth (8/12), and spontaneous abortion (5/0). Among these outcomes, preterm birth had moderate evidence of a positive association with phthalate exposure (specifically DEHP, DBP, and DEP). Exposure levels for BBP, DIBP, and DINP were generally lower than for the phthalates with an observed effect, which may partially explain the difference due to lower sensitivity. Other phthalate/outcome combinations were considered to have slight or indeterminate evidence of an association.

Conclusions and implications of key findings: Overall, these results support that some phthalates may be associated with higher odds of preterm birth in humans, though there is

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

*Corresponding author. radke-farabaugh.elizabeth@epa.gov (E.G. Radke).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.envint.2019.02.003>.

some remaining inconsistency. More evidence is needed on the mechanism and relevant exposure window for this association.

1. Introduction

Phthalates (diesters of phthalic acid) are a class of synthetic chemicals used in many consumer and industrial products, including applications as plasticizers in polyvinyl chloride plastics and as ingredients in some medications and personal care products. Human exposure is ubiquitous across the lifespan, primarily via the oral route, but also through inhalation and dermal contacts (Johns et al., 2015). After exposure, phthalate diesters are rapidly metabolized (estimated half-lives of various phthalate metabolites is approximately 3 to 18 h) to monoester metabolites and excreted in the urine (Johns et al., 2015). The group of phthalates encompasses a variety of compounds with different structures, properties, and use (and relative potency, as shown from animal studies). The most common phthalates, and those focused on here are: di(2-ethylhexyl) phthalate (DEHP), diisononyl phthalate (DINP), dibutyl phthalate (DBP), diisobutyl phthalate (DIBP), butyl benzyl phthalate (BBP), and diethyl phthalate (DEP). The metabolites of each are described in the supplementary materials. Among the group, there are phthalates that are relatively structurally similar and moderately correlated with each other based on human biomonitoring data (e.g., DBP and DIBP). Some phthalates differ considerably in structure and commercial/industrial uses; correlations between these phthalates are typically low.

While male reproductive toxicity is the most commonly discussed hazard of phthalate exposure (NAS, 2017; Johnson et al., 2012; Howdeshell et al., 2008), there has also been substantial study of female reproductive and developmental effects of phthalate exposure. Exposure to some phthalates is higher in women than in men in some studies, which has been hypothesized to be due to higher use of cosmetics (Kay et al., 2013; Fennell et al., 2004; Lovekamp-Swan and Davis, 2003), though this may apply to some phthalates (e.g., DBP, DEP) used in cosmetics more than others. In addition, phthalates can cross the placenta, making developmental effects from in utero exposures a concern (Langonne et al., 1998), particularly given the potential for higher susceptibility in developmental stages. Existing narrative reviews are inconclusive and indicate that there may be an association with some outcomes, most notably preterm birth (Mariana et al., 2016; Hauser and Calafat, 2005; Benjamin et al., 2017), but no systematic review across key female reproductive and developmental effects is available. We performed a systematic review of the epidemiology literature on the female reproductive and developmental effects associated with phthalate exposure. This included review of the following outcomes: pubertal development, time to pregnancy, spontaneous abortion, and preterm birth. In addition, the evidence for some additional outcomes (e.g., ovarian reserve) was reviewed for coherence with the other outcomes and is summarized without systematic review. The human health relevance of these outcomes is summarized briefly in Table 1.

2. Methods

2.1. Literature search and screening

The literature search and screening, study evaluation, data extraction, and evidence synthesis methods are described in detail in the systematic review protocol (Supplement 1, Section 3). Briefly, epidemiology studies were identified by conducting a single broad literature search on all six phthalates of interest (DEHP, DINP, DBP, DIBP, BBP, DEP) and all outcomes. The Population, Exposure, Comparator, and Outcome (PECO) is available in the protocol (Supplement 1, Section 2). The following databases were searched: PubMed, Web of Science, and Toxline, initially in 2013, with updates every 6–12 months through January 2017. Forward and backward searches were also performed. Title/abstract and full text screening was performed by two reviewers to identify studies that met the PECO criteria. In addition, because there was a delay in publication, an updated search and screen was performed in August 2018. The publications from that search that inform each outcome are listed for each outcome and were reviewed to determine if their inclusion would significantly alter the conclusions.

2.2. Study evaluation

Studies were evaluated by at least two reviewers using uniform approaches for each group of similar studies. Key concerns were risk of bias (factors that affect the magnitude or direction of effect) and insensitivity (factors that limit the ability of a study to detect a true effect) (Cooper et al., 2016). Evaluation was conducted for the following domains: exposure measurement, outcome ascertainment, population selection, confounding, analysis, other sensitivity concerns, and selective reporting. These domains were based on the Risk of Bias in Non-randomized Studies of Interventions (ROBINS-I) tool (Sterne et al., 2016), modified for use with environmental exposures by the U.S. Environmental Protection Agency's Integrated Risk Information System. Phthalate and outcome-specific criteria were developed to address issues specific to the evaluation of this chemical and this set of outcomes prior to evaluation. For exposure, most of the available studies relied on phthalate metabolite biomarkers (a list of metabolites for each phthalate is provided in the protocol). Different criteria were developed for short-chain (DEP, DBP, DIBP, BBP) and long-chain (DEHP, DINP) phthalates due to better reliability of single measures for short-chain phthalates. Measurement in urine was considered to be the best proxy of exposure (Johns et al., 2015). Biomarker measures based on samples other than urine (e.g., blood, amniotic fluid, breast milk) were considered to be critically deficient for all short-chain phthalates and for primary metabolites (e.g., MEHP, MINP) of long-chain phthalates. Rationale for these criteria and additional details for all domains are available in the protocol (Supplement 1, Section 4.1.1) and an abbreviated version is available in the key methods supplement (Supplement 2). The outcome-specific criteria are discussed in Radke et al. (forthcoming in special issue).

For each study result, in each evaluation domain, reviewers reached a consensus rating regarding the utility of the study for hazard identification, with categories of *Good*, *Adequate*, *Deficient*, or *Critically deficient*. These ratings were then combined to reach an overall study confidence classification of *High*, *Medium*, *Low*, or *Uninformative*. This overall classification was not based on pre-defined weights for each domain, but rather

was based on reviewer judgments, and include the likely impact the noted deficiencies in bias and sensitivity have on the magnitude and direction of effect estimates, which varies depending on the study and/or outcome. Studies were evaluated for their suitability for each outcome investigated, and could receive different ratings for each outcome. Descriptions of each of the evaluation domains and overall confidence categories can be found in the protocol (Supplement 1, Section 4). Study evaluations were documented in Health Assessment Workspace (HAWC), and ratings and rationale are publicly available (links provided in each outcome section).

2.3. Evidence synthesis

After study evaluation, the evidence for each outcome was synthesized separately for each phthalate, using the following aspects of an association that may suggest causation: consistency, exposure-response relationship, strength of association, temporal relationship, biological plausibility, and coherence. In evaluating the evidence for each of these considerations, syntheses also considered the strengths and limitations of each study, with high confidence studies carrying the most weight. Based on this synthesis, the evidence was assigned a strength of evidence conclusion of *Robust*, *Moderate*, *Slight*, *Indeterminate*, or *Compelling evidence of no effect*. *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* and *Indeterminate* describe evidence for which uncertainties prevents drawing a causal conclusion in either direction. These categories are generally limited in terms of quantity or confidence level of studies, and serve to encourage additional research across the exposure range experienced by humans. *Compelling evidence of no effect* requires several high confidence studies with consistent null results. The ratings for the individual outcomes were then summarized into an overall conclusion for female reproductive and developmental effects by phthalate, using a structured framework available in the key methods supplement (Supplement 2) and the protocol (Supplement 1, Section 6). No statistical quantitative meta-analysis was performed due to substantial differences across studies.

3. Results

The literature flow diagram is shown in Fig. 1 and a list of the included and excluded studies is available at: https://hero.epa.gov/hero/index.cfm/project/page/project_id/2245. For each outcome, there is a summary of the studies included and excluded based on study evaluation, a list of studies identified in the August 2018 update, a presentation of the extracted data (tables and/or figures), and a synthesis of the evidence ending with a judgment of the strength of the evidence in human studies for each phthalate. The studies reviewed from the August 2018 literature update did not influence conclusions, and are not included in Fig. 1.

3.1. Pubertal development

3.1.1. Study selection and evaluation—There were 18 studies (21 papers) that examined female exposure to phthalates and its association with timing of puberty. Thirteen studies were excluded, primarily due to issues with temporality (i.e., exposure was measured post-pubertal onset). Other reasons for exclusion included exposure measurement in tissues

other than urine (unless looking at secondary metabolites of DEHP and DINP), exposure measurement by self-report, or critical deficiency in population selection/study design. The specific phthalate metabolites examined in the remaining five studies (seven papers) and the study evaluations are summarized in Table 2. Full rationales are available at <https://hawcprd.epa.gov/summary/visual/100000090/>.

The included studies varied in sample size, having between 69 and 1170 girls/females included in the analysis. Two studies (Wolff et al., 2014; Mouritsen et al., 2013) examined cohorts of children with phthalate exposure measured prior to pubertal onset. Three studies (Su et al., 2014; Watkins et al., 2014; Hart et al., 2013) were pregnancy cohorts, with exposure measured during gestation and children followed to pubertal onset. The results from the two different time periods (prenatal and childhood) are considered separately in this assessment and both are considered relevant windows of exposure for pubertal onset. For samples collected in utero, the timing in all three studies included the third trimester, with one study (Hart et al., 2013) also collecting samples during the second trimester and one study (Watkins et al., 2017) collecting in all three trimesters. Two studies (Wolff et al., 2014) were classified as medium confidence, and the remaining three (Su et al., 2014; Watkins et al., 2014; Hart et al., 2013; Mouritsen et al., 2013) were classified as low confidence.

3.1.2. Results and synthesis—Evaluation of the association between exposure to DEHP and pubertal development is based on five studies: two with childhood exposure measures (Wolff et al., 2014; Mouritsen et al., 2013) and three with prenatal exposure measures (Su et al., 2014; Watkins et al., 2014; Hart et al., 2013) (Table 3). For childhood exposure, one study reported delayed onset of puberty (Wolff et al., 2014). Results for prenatal exposure were not consistent with that finding, with one study reporting no association (Su et al., 2014; Hart et al., 2013) and two (Watkins et al., 2014; Hart et al., 2013) reporting earlier onset of puberty. Further, the direction of the association was opposite for breast development (delayed with increasing exposure) and pubic hair development (earlier with increasing exposure) in Watkins et al. (2017). Overall, due to lack of consistency and coherence across related measures of puberty, there is *indeterminate* evidence of an effect of DEHP exposure on pubertal development.

Two *low confidence* studies (Mouritsen et al., 2013; Hart et al., 2013) reported on the association between pubertal development and DINP exposure (one with childhood exposure measures, one with prenatal). Neither study reported an association, and this evidence is considered *indeterminate*.

Four studies reported on the association between pubertal development and DBP (two with childhood exposure measures, two with prenatal), although one study reported only a combined estimate for DBP and DIBP (Mouritsen et al., 2013) (Table 4). For childhood exposure, two studies reported later age at pubarche for at least one measure (Wolff et al., 2014; Mouritsen et al., 2013), but results were not consistent across measures. For in utero exposure, one study (Watkins et al., 2017) reported conflicting results for age at menarche (later age with higher MBP exposure) and pubic hair stages (earlier age with higher exposure), the latter of which also conflicted with the results for childhood exposure.

In Su et al. (2014), no association was reported. Overall, because of inconsistency and lack of coherence between related measures of puberty, there is *indeterminate* evidence of an effect of DBP exposure on pubertal development.

There are two studies that reported on the potential association of DIBP exposure with pubertal development (one with childhood exposure measures, one with prenatal). For prenatal exposure, Watkins et al. (2014) reported results for menarche and breast and pubic hair development in opposite directions. Other results indicated no association. Overall, there is *indeterminate* evidence of an effect of DIBP exposure on pubertal development.

Evaluation of the evidence of an association between BBP exposure and pubertal development is based on four studies (two with childhood exposure measures, two with prenatal) (Table 5). For childhood exposure, the studies had conflicting results, with one showing earlier onset of puberty (Mouritsen et al., 2013) for breast development and one showing delayed onset of puberty for breast and pubic hair development (Wolff et al., 2014). The effect sizes were small. For prenatal exposure, Watkins et al. (2014) reported earlier onset of breast and pubic hair development, but later onset of menarche with increasing exposure, while Su et al. (2014) reported no association. Overall, there is *indeterminate* evidence of an effect of BBP exposure on pubertal development.

Four studies reported on the association between pubertal development and DEP (two with childhood exposure measures, two with prenatal) (Table 5). For childhood exposure, one study reported earlier onset of puberty for age at menarche (Wolff et al., 2017) and the other study reported no association (Mouritsen et al., 2013). For prenatal exposure, one study (Watkins et al., 2014) showed earlier onset at pubarche with increasing exposure, while the other study (Su et al., 2014) reported no association. Because earlier onset of puberty was observed in both medium confidence studies, overall, there is *slight* evidence of an effect of DEP exposure on pubertal development.

In summary, there is *indeterminate* evidence of an association between pubertal development in females and exposure to phthalates, with the exception of DEP, which has *slight* evidence. Direct comparison of the effect estimates across studies was hindered by the different analysis methods and units, thus the synthesis focused on direction of association, which was inconsistent across endpoints within studies, across studies, and across phthalates, including in medium confidence studies.

3.2. Fecundity

3.2.1. Study selection and evaluation—There were seven studies that examined female exposure to phthalates and its association with time to pregnancy, and three were included in the evaluation. Four studies were excluded (Table 6), primarily due to issues with temporality (i.e., exposure after outcome). The specific phthalate metabolites examined in the remaining three studies and the study evaluations are summarized in Table 6. One additional study examined another outcome related to fecundity, rates of clinical pregnancy (Hauser et al., 2016). Full rationale is available at <https://hawcprd.epa.gov/summary/visual/100000091/>.

All the included studies of time to pregnancy were prospective cohorts of couples trying to conceive. Time to pregnancy was determined prospectively. All studies employed Cox regression methods to calculate fecundability ORs, in which a lower value (< 1) represents reduced fecundity (increased time to pregnancy). Buck Louis et al. (2014) collected only one urine sample for exposure measurement, while the other two studies analyzed multiple samples when available. Thomsen et al. (2017) included only 242 of the 430 women in the original cohort due to missing urine samples (not collected due to conception in the first cycle or lost during storage). Because time to pregnancy differed in those without urine samples, this was considered a possible selection bias. One study was classified as *high confidence* (Jukic et al., 2016) and two studies were classified as *medium confidence* (Thomsen et al., 2017; Buck Louis et al., 2014). The single study that evaluated rates of pregnancy (Hauser et al., 2016), was a cohort of women undergoing in vitro fertilization and was classified as *high confidence*.

3.2.2. Results and synthesis—Three studies examined the association between DEHP exposure in women and time to pregnancy. Results for metabolites MEHP and MEOHP are provided in Table 7. There is no evidence that higher DEHP exposure is associated with longer time to pregnancy. However, exposure levels were fairly low across the studies, which may have reduced the ability of individual studies to detect associations. The one study examining rates of clinical pregnancy in a population of couples seeking fertility treatment (Hauser et al., 2016) reported lower percentages of participants with a pregnancy with higher exposure to DEHP (Q1: 0.57, 95% CI = 0.45–0.69, Q2: 0.46, 95% CI = 0.36–0.57; Q3: 0.49, 95% CI = 0.38–0.59; Q4: 0.38, 95% CI = 0.28–0.49*, p-trend = 0.04). In addition, three studies (four publications) examined related outcomes in women undergoing in vitro fertilization, and reported decreases in oocytes (Hauser et al., 2016; Machtinger et al., 2018), antral follicle count (Messerlian et al., 2015), embryo quality (Machtinger et al., 2018), and implantation (Hauser et al., 2016), though one study reported no decrease in embryo quality (Wu et al., 2017). Still, because of the lack of supporting evidence from studies on time to pregnancy, which were all of *high* or *medium confidence*, evidence of an association between DEHP exposure in women and fecundity is considered *slight*.

Evaluation of the association between exposure to the other phthalates and time to pregnancy is based on a subset of the studies described for DEHP, with one to three studies for each phthalate (Table 8). With the exception of one study for DEP (Thomsen et al., 2017), no association was reported between higher phthalate exposure and lower fecundity/increased time to pregnancy. This included the single study examining rates of clinical pregnancy (Hauser et al., 2016). There were decreases in related outcomes in women undergoing in vitro fertilization in three studies (four publications) for DBP, DIBP, and DEP in at least one secondary outcome (Hauser et al., 2016; Messerlian et al., 2015; Machtinger et al., 2018; Wu et al., 2017), but based on the lack of association for the primary fecundity outcomes, evidence of an association between fecundity and exposure to DINP, DBP, DIBP, BBP, and DEP is considered *indeterminate*.

3.3. Spontaneous abortion

3.3.1. Study selection and evaluation—There were five studies that examined exposure to phthalates and its association with spontaneous abortion; none were excluded. The specific phthalate metabolites examined in the studies and the study evaluations are summarized in Table 9. Full rationale is available at <https://hawcprd.epa.gov/summary/visual/100000092/>.

Two studies were preconception general population cohorts (Jukic et al., 2016; Toft et al., 2012), one was a preconception cohort among women receiving assisted reproductive technology (Messerlian et al., 2016) and two were clinic-based, case-control studies (Mu et al., 2015); (Yi et al., 2016). With regards to population selection, the case-control studies only included women who received ultrasound to confirm pregnancy, which represented a potential selection bias as it excludes women who experienced a spontaneous abortion and did not receive an ultrasound. One of the preconception cohorts (Toft et al., 2012) had urine samples for only 56% (242 out of 430) of the cohort participants because urine samples were not collected if conception occurred during the first cycle ($n = 111$) or because urine samples were lost during the storage period. Time to pregnancy differed between the participants with and without urine samples, which was considered a potential source of selection bias. Two studies were classified as having *high confidence* (Jukic et al., 2016; Messerlian et al., 2016) and the other three (Mu et al., 2015; Toft et al., 2012); (Yi et al., 2016) were classified as having *low confidence*.

3.3.2. Results and synthesis—Evaluation of the evidence of an association between exposure to DEHP and spontaneous abortion (Table 10) is based on five studies, with three studies reporting on early loss, three studies reporting on clinical loss, and one reporting on total loss. Of the two high confidence studies, Messerlian et al. (2016) reported elevated odds of total loss (early and clinical loss combined) with higher exposure while Jukic et al. (2016) reported lower odds of early loss with higher exposure. For early and clinical loss, two low confidence studies reported elevated odds with higher exposure (Toft et al. (2012) for early loss and Yi et al. (2016) for clinical loss), though the effect estimates for one (Mu et al., 2015; Toft et al., 2012) were imprecise. Overall, there is *slight* evidence of an association between DEHP exposure and spontaneous abortion, but there is considerable uncertainty due to inconsistency in the *high* confidence studies.

DINP was measured in two studies (Jukic et al., 2016; Messerlian et al., 2016). No association with early loss or total loss, respectively, was observed. This evidence is considered *indeterminate*.

Evaluation of the evidence of an association between exposure to DBP and spontaneous abortion is based on five studies (Table 11). Jukic et al. (2016), a *high confidence* study, found slightly elevated ORs between early loss and MBP exposure levels. Toft et al. (2012) observed a monotonic increase in OR for early loss with increasing exposure, and an inverse association for clinical loss with increasing exposure. Mu et al. (2015) reported an elevated OR for clinical loss for quartile 4 vs. quartile 1, but an inverse association for quartiles 2 and 3 vs. quartile 1. None of the reported associations were statistically significant. Messerlian et al. (2016), another *high confidence* study, and Yi et al. (2016) reported no association. The

results for clinical loss were inconsistent, and the effect estimates for early loss were small and not statistically significant. Overall, there is *slight* evidence of an association between DBP exposure and early spontaneous abortion.

Four studies examined the association between spontaneous abortion and DIBP exposure (Table 11). The two *high confidence* studies, Jukic et al. (2016) and Messerlian et al. (2016) found no relationship between MiBP and early pregnancy loss and total pregnancy loss; however, DIBP exposure levels were lower than those reported for other phthalates and this may have affected study sensitivity. One low confidence study (Yi et al., 2016) also found no increase in clinical loss. Conversely, Mu et al. (2015) reported generally increasing associations with increasing exposure, including a statistically significant finding in the 4th vs. 1st quartile. Given the lack of consistency across studies and the reporting of an association in only a single *low confidence* study, the evidence of an association between DIBP exposure and spontaneous abortion is considered *indeterminate*.

Three studies examined the evidence of an association between BBP exposure (measured by MBzP) and spontaneous abortion (Table 11). Toft et al. (2012) found higher ORs with a monotonic trend for early loss, but no association between exposure and clinical loss. Results from Jukic et al. (2016) and Messerlian et al. (2016), *high confidence* studies, did not support this finding despite Jukic et al. (2016) reporting higher BBP exposure levels. Based on this lack of consistency across studies, the evidence of an association between BBP exposure and spontaneous abortion is considered *indeterminate*.

Five studies examined the evidence of an association between DEP exposure and spontaneous abortion. Jukic et al. (2016) and Messerlian et al. (2016), the *high confidence* studies, found no association between MEP levels and early loss or total loss, respectively (Table 11). One *low confidence* study (Yi et al., 2016) also reported no association. Toft et al. (2012) reported slightly elevated ORs for early loss, and moderately elevated ORs for clinical loss. Mu et al. (2015) reported a monotonic increase in ORs with increasing exposure. There was not a clear explanation for the lack of association in the *high confidence* studies, but it is possible that the association is limited to clinical loss. Given this lack of consistency, the evidence for an association between DEP exposure and spontaneous abortion is *slight*.

In summary, there is *slight* evidence of an association between spontaneous abortion and exposure to DEHP, DBP, and DEP. There is generally inconsistency among the high confidence studies.

3.4. Preterm birth

3.4.1. Study selection and evaluation—There were 20 studies (described in 22 publications) that examined exposure to phthalates and its association with preterm birth or gestational duration. Four studies were excluded due to issues with exposure measurement and six studies were excluded due to participant enrollment late in pregnancy (i.e., a majority of participants were enrolled during the third trimester) (Table 12). The specific phthalate metabolites examined in the remaining eight studies (ten publications) and the

study evaluations are summarized in Table 12 and full rationale is available at <https://hawcprd.epa.gov/summary/visual/100000093/>.

Six studies are pregnancy cohorts, and two are case-control studies nested within pregnancy cohorts. These studies vary in size between 60 and 3103 mother-infant pairs included in the analysis. Four studies (Casas et al., 2016; Shoaff et al., 2016; Ferguson et al., 2014b) measured exposure in two or more urine samples, with the levels averaged, while the remaining studies used a single urine sample. Only Ferguson et al. (2014b) additionally analyzed the exposure collections separately. The timing of collection (preconception, 1st, 2nd, or 3rd trimester) varied among the studies; timing variability during pregnancy was not considered to be a basis for downgrading confidence in the results because the critical window is not known. However, preconception exposure resulted in a downgrade in confidence because women may change their diet and overall exposure when they become pregnant, and pregnancy may cause metabolic changes. Five studies (Casas et al., 2016; Gao et al., 2016; Pola ska et al., 2016; Watkins et al., 2016; Smarr et al., 2015) analyzed gestational duration as a continuous variable; however, authors of three of these studies (Casas et al., 2016; Gao et al., 2016; Smarr et al., 2015) were contacted and were able to provide analyses with preterm birth as the outcome to improve comparability between studies, and those results are presented here rather than the gestational duration results. Only one study (Ferguson et al., 2014b) stratified by type of preterm birth (i.e., spontaneous preterm birth). Three studies (Casas et al., 2016; Shoaff et al., 2016; Ferguson et al., 2014b) were classified as *high confidence*, and the remaining five were classified as *medium confidence*. The bias for all studies was considered likely to be towards the null, if present, due to the possibility of exposure misclassification (see exposure criteria in Supplement 1, Section 4.1.1) and other study-specific limitations.

3.4.2. Results and synthesis—Evaluation of the evidence of an association between exposure to DEHP and preterm birth is based on six studies that analyzed preterm birth as a dichotomous variable (Casas et al., 2016; Gao et al., 2016; Shoaff et al., 2016; Smarr et al., 2015; Ferguson et al., 2014b; Meeker et al., 2009) (Fig. 2, results are sorted by mean exposure level for studies with continuous exposure) and two studies that analyzed gestational duration as a continuous variable (Pola ska et al., 2016; Watkins et al., 2016). Of the preterm birth studies, four of the six, including two high confidence studies, reported elevated ORs, with one (Ferguson et al., 2014b) reaching statistical significance, and very similar results among these four studies (OR range for continuous exposure: 1.25–1.33). An exposure-response relationship was observed in Ferguson et al. (2014b) (OR range: 1.23–2.17) and Casas et al. (2016), although the relationship was non-monotonic in the latter study. The largest study (Gao et al., 2016) reported no association, but also had lower levels of exposure than other studies, with only Smarr et al. (2015) reporting lower levels of MEOHP (not shown). This may indicate reduced study sensitivity. The two gestational duration studies reported no association between pregnancy duration and DEHP exposure, but this is not considered inconsistent with the preterm birth results due to the limitations in analyzing gestational duration as a proxy for preterm birth. While there is consistency for preterm birth among multiple medium and high confidence studies in varied settings (e.g., multiple different countries), there is a lack of association in one high confidence study, as

well as the largest study, which may be partly explained by lower study sensitivity resulting from low exposure levels. Therefore, there is *moderate* evidence of an association between DEHP exposure and preterm birth.

Two studies of *medium confidence* included results for DINP. Meeker et al. (2009) reported slightly elevated odds (not statistically significant) of preterm birth with higher DINP exposure in a small study (MCiOP OR = 1.3, 95% CI = 0.5–3.9), while Pola ska et al., 2016 reported no association with gestational duration (DINP $\beta = -0.03$, SE = 0.08, $p = 0.5$). The evidence for an association between DINP exposure and preterm birth from this one study is considered *slight*.

Evaluation of the evidence of an association between exposure to DBP and preterm birth is based on six studies (Casas et al., 2016; Gao et al., 2016; Shoaff et al., 2016; Smarr et al., 2015; Ferguson et al., 2014b; Meeker et al., 2009) that analyzed preterm birth as a dichotomous variable (Fig. 3; results are sorted by mean exposure level for studies with continuous exposure) and two studies that analyzed gestational duration as a continuous variable (Pola ska et al., 2016; Watkins et al., 2016). Three studies, including two high confidence, reported elevated odds of preterm birth (OR range: 1.3 to 4.5) with higher exposure (Casas et al., 2016; Ferguson et al., 2014b; Meeker et al., 2009). Statistically significant results were reported in Ferguson et al. (2014b) and Meeker et al. (2009), and the OR in Meeker et al. (2009) was high. The remaining three studies (Gao et al., 2016; Shoaff et al., 2016; Smarr et al., 2015) reported no association or an association with longer gestation; however, the DBP exposure levels in Smarr et al. (2015) were the lowest of the studies, and as such the study may have had lower sensitivity to detect an association. Shoaff et al. (2016) reported statistically significant lower odds of preterm birth with higher exposure. Among the two gestational duration studies, the only association reported between pregnancy duration and DBP exposure was in female infants in Watkins et al., 2016 (for male infants change in days = 0.09, 95% CI = -3.81, 3.99; for female infants change in days = -4.46, 95% CI = -8.25, -0.67*). Overall, while there are some inconsistencies in the results across studies, there is *moderate* evidence of an association between DBP exposure and preterm birth.

Five studies on preterm birth (Casas et al., 2016; Gao et al., 2016; Shoaff et al., 2016; Smarr et al., 2015; Ferguson et al., 2014b; Meeker et al., 2009) and one study on gestational duration (Pola ska et al., 2016) investigated the association with exposure to DIBP (Fig. 3). One *high confidence* study (Shoaff et al., 2016) and one *medium confidence* study (Meeker et al., 2009) reported non-statistically significant elevated odds of preterm birth with higher DIBP exposure, while the other three studies reported no association between exposure and preterm birth, and there was no clear pattern between exposure level or range and higher ORs. Pola ska et al. (2016) reported no association with gestational duration ($\beta = 0.01$, SE = 0.04, $p = 0.8$). DIBP exposure levels were low in all of the studies relative to other phthalates, with the exception of Casas et al. (2016), so the sensitivity to detect an association may have been inadequate. The current evidence is considered *slight*.

Evaluation of the evidence for an association between exposure to BBP and preterm birth is based on six studies (Casas et al., 2016; Gao et al., 2016; Shoaff et al., 2016; Smarr

et al., 2015; Ferguson et al., 2014b; Meeker et al., 2009) that analyzed preterm birth as a dichotomous variable (Fig. 3; results are sorted by mean exposure level for studies with continuous exposure variables) and one study that analyzed gestational duration as a continuous variable (Watkins et al., 2016). One *high confidence* study Ferguson et al. (2014b) and one *medium confidence* study (Meeker et al., 2009) reported non-statistically significant elevated odds of preterm birth with increasing BBP exposure, although in Ferguson et al. (2014b) the OR was only slightly elevated. Results from the other four studies (two *high confidence* and two *medium confidence* studies) were either no association or lowered odds. There was no increase in ORs with increasing exposure levels or range. The gestational duration study reported no association between pregnancy duration and BBP exposure (Watkins et al. 2016) male infants change in days ($\beta = 0.27$, 95% CI = -4.35, 4.90; female infants change in days ($\beta = -2.60$, 95% CI = -7.87, 2.67)). However, BBP exposure in all of the studies was low relative to other phthalates, which may have reduced the sensitivity to detect an association. Therefore, the evidence is considered *slight*.

Of the same six studies that investigated DEP exposure and preterm birth, four studies, including all three high confidence studies (Casas et al., 2016; Shoaff et al., 2016; Ferguson et al., 2014b; Meeker et al., 2009) reported consistent higher odds (OR range: 1.1–1.4) of preterm birth with increasing exposure. These results were not statistically significant, and the effect estimates were smaller than the results for other phthalates, particularly in Meeker et al. (2009). Conversely, Smarr et al. (2015) reported a statistically significant negative association and Gao et al. (2016) reported no association, but these two studies had the lowest exposure levels. Among the two gestational duration studies, both reported an association between shorter gestational duration and increasing exposure, with statistical significance in one (Pola ska et al. (2016) $\beta = -0.2$, SE = 0.1, $p = 0.04^*$; Watkins et al. (2016) for male infants change in days = -3.30, 95% CI = -7.94, 1.34; for female infants change in days = -2.86, 95% CI = -9.19, 3.47). Given the reasonable consistency between the studies, the evidence that DEP exposure is associated with higher risk of preterm birth is considered *moderate*.

In addition to the primary effect, one study (Ferguson et al., 2016) showed that oxidative stress may act as a mediator between phthalate exposure and preterm birth, based on a strong association between elevated urinary phthalate metabolites and higher levels of oxidative stress biomarkers in pregnant women. The author noted that oxidative stress can cause apoptosis at the maternal-fetal interface, leading to poor vascularization of the placenta and consequently preeclampsia and/or intrauterine growth restriction during pregnancy. Another study (Adibi et al., 2010) reported that increased urinary phthalate metabolite concentrations were associated with a significant downregulation in the expression of placental genes involved in trophoblast differentiation (PPAR γ , AhR, and HCG). There is also evidence from animal studies that phthalates interfere with maternal ovarian function and steroidogenesis, which may have effects on pregnancy maintenance; however, this mechanism in females remains poorly defined (reviewed in Kay et al., 2013). This mechanistic support provides some biological plausibility for the observed associations.

In summary, there is *moderate* evidence of an association between preterm birth and exposure to DEHP, DBP, and DEP. Table 13 presents a summary of the factors that increase

and decrease confidence in this evidence. The evidence of an association for the remaining phthalates is considered *slight*. Exposure levels for BBP, DIBP, and DINP were generally lower than for the phthalates with an observed association, which may at least partially explain the differences. On the other hand, in the study where DBP and DIBP levels were similar Casas et al. (2016), there was an association observed for DBP (not statistically significant) but not DIBP.

4. Discussion

Of the outcomes reviewed, preterm birth has the strongest evidence of an association with phthalate exposures (Fig. 4). Three of the six phthalates had *moderate* evidence of a positive association, and the remaining three had *slight* evidence but also fewer studies and/or lower exposure levels that could explain the discrepancy. Looking across, there is evidence that preterm birth is a hazard associated with at least some phthalate exposures in humans. A 1-unit increase in the level of phthalate metabolites was associated with a 20–30% or greater increase in the odds of preterm birth in some studies. This level of increase is of clear public health concern.

Looking across the other outcomes, there is little convincing evidence of an association between phthalate exposure and female reproductive outcomes. Spontaneous abortion had *slight* evidence of an association for some phthalates. More studies on spontaneous abortion with repeated exposure measures, high study sensitivity (e.g., varied exposure levels, large sample size), and prospective outcome ascertainment would help clarify the association. This is particularly true for DEHP since the results of one high confidence study were compelling but the low confidence studies were conflicting. For pubertal development and time to pregnancy, the findings were almost all *indeterminate*, but additional *high* confidence studies may be informative here as well.

For some outcome-phthalate pairs, lower exposure levels than for other phthalates decreased the sensitivity to observe associations with health effects even if they were present in the population and hindered the interpretability of null findings. This reduced the ability to assess coherence across phthalates. It also is difficult to evaluate the shape of the exposure-response curve in groups exposed at relatively low levels. While there is evidence of non-monotonic dose-response relations between some endocrine disrupting chemicals and health outcomes (Vandenberg et al., 2012), such evidence does not exist for female reproductive outcomes and studies of male reproductive outcomes indicate that such relations are linear (NAS, 2017). It seems unlikely that we would be able to precisely identify any non-monotonic associations given that phthalate biomarker concentrations typically span 2–4 orders of magnitude in most epidemiological studies and human phthalate exposures are several orders of magnitude lower than those used in animal toxicology studies. However, future studies should employ appropriate methods (e.g., smoothing splines) to identify the presence of any non-monotonic dose-response relations.

Overall, these findings are mostly consistent with past narrative reviews. Mariana et al. (2016) found that the evidence does not support an association between phthalate exposure and precocious puberty and reported weak evidence of an association with birth

size. Further, they found that there may be an association between prenatal exposure to phthalates and reduced gestational time. Benjamin et al. (2017) reported associations between phthalates and low gestational duration and spontaneous abortion but did not provide a critical review of the available literature. Both of these reviews included the same compounds examined in this review.

This systematic review benefited from the involvement of experts on phthalates and female reproductive and developmental epidemiology throughout the process, from development of the study evaluation criteria through evidence synthesis. Another strength is that methods for study evaluation and evidence synthesis were predefined in the systematic review protocol, and that studies were evaluated by two reviewers to reduce potential for bias. While study evaluation and evidence synthesis conclusions relied on expert judgment of the reviewers using structured frameworks, this is preferred in the systematic review community due to the limitations of more prescriptive, score-based approaches that tend to focus on reporting quality rather than the impact of identified limitations on the results (Higgins and Green, 2011; Herbison et al., 2006; Juni et al., 1999).

However, while systematic review tools have been in use for some time in studies of clinical interventions, these methods have been introduced in the field of environmental health assessment more recently. Modifications to the existing tools and approaches have been necessary to accommodate challenges inherent to observational studies where exposure is not controlled. However, since the application of these approaches is newer and less tested, we recognize that there may be aspects of the methodology that could benefit from greater transparency for the reader, such as communication of inputs to decisions.

There are other important limitations to note. Due to remaining uncertainty in the database, including inconsistency among some high confidence studies and potential for remaining unidentified confounders, it is not possible to conclusively establish causality from the observational epidemiology studies alone. The integration of this evidence with animal toxicology studies and mechanistic studies could provide additional information to address some of the uncertainty in the epidemiologic evidence and increase the strength of the causal analysis. A complete mechanistic analysis would require additional systematic literature identification and synthesis and was outside the scope of this review. In addition, while a quantitative meta-analysis may have been informative for studies of preterm birth, it was not performed because the necessary analyses were not provided in the same way by all of the studies. A meta-analysis on this topic would ideally be performed using the original datasets, to allow for better exploration of the exposure-response pattern across the studies that contribute data from different exposure ranges, however, this was outside the scope of this review. Discussion of some additional issues that apply to all phthalate studies is presented in a forthcoming editorial. These include the difficulties in assessing health effects of mixtures and separating individual contributions from highly correlated phthalates. Despite these limitations, our systematic review of the epidemiological literature supports an association between some phthalate exposures and preterm birth.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to thank Erin Yost, Rebecca Nachman, Susan Rieth, Kris Thayer, Mike Wright, Emily Harville, Anna Pollack, and Andrew Greenhalgh for their contributions to this work.

References

- Adibi JJ, Hauser R, Williams PL, Whyatt RM, Calafat AM, Nelson H, Herrick R, Swan SH, 2009. Maternal urinary metabolites of di-(2-ethylhexyl) phthalate in relation to the timing of labor in a US multicenter pregnancy cohort study. *Am. J. Epidemiol.* 169, 1015–1024. 10.1093/aje/kwp001. [PubMed: 19251754]
- Adibi JJ, Whyatt RM, Hauser R, Bhat HK, Davis BJ, Calafat AM, Hoepner LA, Perera FP, Tang D, Williams PL, 2010. Transcriptional biomarkers of steroidogenesis and trophoblast differentiation in the placenta in relation to prenatal phthalate exposure. *Environ. Health Perspect.* 118, 291–296. 10.1289/ehp.0900788. [PubMed: 20123604]
- Benjamin S, Masai E, Kamimura N, Takahashi K, Anderson RC, Faisal PA, 2017. Phthalates impact human health: epidemiological evidences and plausible mechanism of action [review]. *J. Hazard. Mater.* 340, 360–383. 10.1016/j.jhazmat.2017.06.036. [PubMed: 28800814]
- Binder AM, Corvalan C, Calafat AM, Ye X, Mericq V, Pereira A, Michels KB, 2018. Childhood and adolescent phenol and phthalate exposure and the age of menarche in Latina girls. *Environ. Health Perspect.* 126, 118–124. 10.1289/ehp.2017.126.118. [PubMed: 29615064]
- Brucker-Davis F, Wagner-Mahler K, Bornebusch L, Delattre I, Ferrari P, Gal J, Boda-Buccino M, Pacini P, Tommasi C, Azuar P, Bongain A, Fénelon P, 2010. Exposure to selected endocrine disruptors and neonatal outcome of 86 healthy boys from Nice area (France). *Chemosphere* 81, 169–176. 10.1016/j.chemosphere.2010.06.068. [PubMed: 20663538]
- Buck Louis GM, Schisterman EF, Sweeney AM, Wilcosky TC, Gore-Langton RE, Lynch CD, Boyd Barr D, Schrader SM, Kim S, Chen Z, Sundaram R, 2011. Designing prospective cohort studies for assessing reproductive and developmental toxicity during sensitive windows of human reproduction and development—the LIFE study. *Paediatr. Perinat. Epidemiol.* 25, 413–424. 10.1111/j.1365-3016.2011.01205.x. [PubMed: 21819423]
- Buck Louis GM, Sundaram R, Sweeney AM, Schisterman EF, Maisog J, Kannan K, 2014. Urinary bisphenol A, phthalates, and couple fecundity: the Longitudinal Investigation of Fertility and the Environment (LIFE) study. *Fertil. Steril.* 101, 1359–1366. 10.1016/j.fertnstert.2014.01.022. [PubMed: 24534276]
- Buck Louis GM, Smarr MM, Sun L, Chen Z, Honda M, Wang W, Karthikraj R, Weck J, Kannan K, 2018. Endocrine disrupting chemicals in seminal plasma and couple fecundity. *Environ. Res.* 163, 64–70. 10.1016/j.envres.2018.01.028. [PubMed: 29426029]
- Bulu AD, Aci A, Erkekoglu P, Balci A, Andiran N, Koçer-Gümü el B, 2016. The evaluation of possible role of endocrine disruptors in central and peripheral precocious puberty. *Toxicol. Mech. Methods* 26, 493–500. 10.3109/15376516.2016.1158894. [PubMed: 27451808]
- Burdorf A, Brand T, Jaddoe VW, Hofman A, Mackenbach JP, Steegers EA, 2011. The effects of work-related maternal risk factors on time to pregnancy, preterm birth and birth weight: the Generation R Study. *Occup. Environ. Med.* 68, 197–204. 10.1136/oem.2009.046516. [PubMed: 21172792]
- Buttke DE, Sircar K, Martin C, 2012. Exposures to endocrine-disrupting chemicals and age of menarche in adolescent girls in NHANES (2003–2008). *Environ. Health Perspect.* 120, 1613–1618. 10.1289/ehp.1104748. [PubMed: 23124194]
- Carran M, Shaw IC, 2012. New Zealand Malayan war veterans' exposure to dibutylphthalate is associated with an increased incidence of cryptorchidism, hypospadias and breast cancer in their children. *N. Z. Med. J.* 125, 52–63.

- Casas M, Valvi D, Ballesteros-Gomez A, Gascon M, Fernández MF, Garcia-Esteban R, Iñiguez C, Martínez D, Murcia M, Monfort N, Luque N, Rubio S, Ventura R, Sunyer J, Vrijheid M, 2016. Exposure to bisphenol A and phthalates during pregnancy and ultrasound measures of fetal growth in the INMA-Sabadell cohort. *Environ. Health Perspect.* 124, 521–528. 10.1289/ehp.1409190. [PubMed: 26196298]
- Chen CY, Chou YY, Wu YM, Lin CC, Lin SJ, Lee CC, 2013. Phthalates may promote female puberty by increasing kisspeptin activity. *Hum. Reprod.* 28, 2765–2773. 10.1093/humrep/det325. [PubMed: 23945596]
- Chen YH, Ferguson KK, Meeker JD, Mcelrath TF, Mukherjee B, 2015. Statistical methods for modeling repeated measures of maternal environmental exposure biomarkers during pregnancy in association with preterm birth. *Environ. Health* 14, 9. 10.1186/1476-069X-14-9. [PubMed: 25619201]
- Chou YY, Huang PC, Lee CC, Wu MH, Lin SJ, 2009. Phthalate exposure in girls during early puberty. *J. Pediatr. Endocrinol. Metab.* 22, 69–77. [PubMed: 19344077]
- Cooper G, Lunn R, Agerstrand M, Glenn B, Kraft A, Luke A, Ratcliffe J, 2016. Study sensitivity: evaluating the ability to detect effects in systematic reviews of chemical exposures. *Environ. Int.* 92–93, 605–610. 10.1016/j.envint.2016.03.017.
- Fennell TR, Krol WL, Sumner SCJ, Snyder RW, 2004. Pharmacokinetics of dibutylphthalate in pregnant rats. *Toxicol. Sci.* 82, 407–418. 10.1093/toxsci/kfh294. [PubMed: 15456918]
- Ferguson KK, Mcelrath TF, Ko YA, Mukherjee B, Meeker JD, 2014a. Variability in urinary phthalate metabolite levels across pregnancy and sensitive windows of exposure for the risk of preterm birth. *Environ. Int.* 70C, 118–124. 10.1016/j.envint.2014.05.016.
- Ferguson KK, Mcelrath TF, Meeker JD, 2014b. Environmental phthalate exposure and preterm birth. *JAMA Pediatr.* 168, 61–67. 10.1001/jamapediatrics.2013.3699. [PubMed: 24247736]
- Ferguson KK, Chen YH, Vanderweele TJ, Mcelrath TF, Meeker JD, Mukherjee B, 2016. Mediation of the relationship between maternal phthalate exposure and preterm birth by oxidative stress with repeated measurements across pregnancy. *Environ. Health Perspect.* 125, 488–494. 10.1289/EHP282. [PubMed: 27352406]
- Frederiksen H, Sørensen K, Mouritsen A, Aksglaede L, Hagen CP, Petersen JH, Skakkebaek NE, Andersson AM, Juul A, 2012. High urinary phthalate concentration associated with delayed pubarche in girls. *Int. J. Androl.* 35, 216–226. 10.1111/j.1365-2605.2012.01260.x. [PubMed: 22428786]
- Gao H, Xu YY, Huang K, Ge X, Zhang YW, Yao HY, Xu YQ, Yan SQ, Jin ZX, Sheng J, Zhu P, Hao JH, Tao FB, 2016. Cumulative risk assessment of phthalates associated with birth outcomes in pregnant Chinese women: a prospective cohort study. *Environ. Pollut.* 222, 549–556. 10.1016/j.envpol.2016.11.026. [PubMed: 28024814]
- Gao H, Zhang YW, Huang K, Yan SQ, Mao LJ, Ge X, Xu YQ, Xu YY, Sheng J, Jin ZX, Zhu P, Tao XG, Hao JH, Tao FB, 2017. Urinary concentrations of phthalate metabolites in early pregnancy associated with clinical pregnancy loss in Chinese women: supplementary materials [supplemental data]. *Sci. Rep.* 7.
- Gnoth C, Godehardt D, Godehardt E, Frank-Herrmann P, Freundl G, 2003. Time to pregnancy: results of the German prospective study and impact on the management of infertility. *Hum. Reprod.* 18, 1959–1966. [PubMed: 12923157]
- Hart R, Doherty DA, Frederiksen H, Keelan JA, Hickey M, Sloboda D, Pennell CE, Newnham JP, Skakkebaek NE, Main KM, 2013. The influence of antenatal exposure to phthalates on subsequent female reproductive development in adolescence: a pilot study. *Reproduction* 147, 379–390. 10.1530/REP-130331.
- Hashemipour M, Kelishadi R, Amin MM, Ebrahim K, 2018. Is there any association between phthalate exposure and precocious puberty in girls? *Environ. Sci. Pollut. Res. Int.* 25, 13589–13596. 10.1007/s11356-018-1567-4. [PubMed: 29497942]
- Hauser R, Calafat AM, 2005. Phthalates and human health [review]. *Occup. Environ. Med.* 62, 818. 10.1136/oem.2004.017590.
- Hauser R, Gaskins AJ, Souter I, Smith KW, Dodge LE, Ehrlich S, Meeker JD, Calafat AM, Williams PL, 2016. Urinary phthalate metabolite concentrations and reproductive outcomes among women

- undergoing in vitro fertilization: results from the EARTH study. *Environ. Health Perspect.* 124, 831–839. 10.1289/ehp.1509760. [PubMed: 26545148]
- Herbison P, Hay-Smith J, Gillespie WJ, 2006. Adjustment of meta-analyses on the basis of quality scores should be abandoned [review]. *J. Clin. Epidemiol.* 59, 1249–1256. 10.1016/j.jclinepi.2006.03.008. [PubMed: 17098567]
- Higgins JPT, Green S, 2011. *Cochrane Handbook for Systematic Reviews of Interventions*, Version 5.1.0 (Updated March 2011). The Cochrane Collaboration Retrieved from. <http://handbook.cochrane.org/>.
- Howdeshell KL, Rider CV, Wilson VS, Gray LE Jr., 2008. Mechanisms of action of phthalate esters, individually and in combination, to induce abnormal reproductive development in male laboratory rats [review]. *Environ. Res.* 108, 168–176. 10.1016/j.envres.2008.08.009. [PubMed: 18949836]
- Huang PC, Kuo PL, Chou YY, Lin SJ, Lee CC, 2009. Association between prenatal exposure to phthalates and the health of newborns. *Environ. Int.* 35, 14–20. 10.1016/j.envint.2008.05.012. [PubMed: 18640725]
- Huang Y, Li J, Garcia JM, Lin H, Wang Y, Yan P, Wang L, Tan Y, Luo J, Qiu Z, Chen JA, Shu W, 2014. Phthalate levels in cord blood are associated with preterm delivery and fetal growth parameters in Chinese women. *PLoS ONE* 9, e87430. 10.1371/journal.pone.0087430.
- Huang PC, Tsai CH, Liang WY, Li SS, Huang HB, Kuo PL, 2016. Early phthalates exposure in pregnant women is associated with alteration of thyroid hormones. *PLoS ONE* 11, e0159398. 10.1371/journal.pone.0159398.
- Johns LE, Cooper GS, Galizia A, Meeker JD, 2015. Exposure assessment issues in epidemiology studies of phthalates [review]. *Environ. Int.* 85, 27–39. 10.1016/j.envint.2015.08.005. [PubMed: 26313703]
- Johnson KJ, Heger NE, Boekelheide K, 2012. Of mice and men (and rats): phthalate-induced fetal testis endocrine disruption is species-dependent [review]. *Toxicol. Sci.* 129, 235–248. 10.1093/toxsci/kfs206. [PubMed: 22700540]
- Jukic AM, Calafat AM, McConaughy DR, Longnecker MP, Hoppin JA, Weinberg CR, Wilcox AJ, Baird DD, 2016. Urinary concentrations of phthalate metabolites and bisphenol A and associations with follicular-phase length, luteal-phase length, fecundability, and early pregnancy loss. *Environ. Health Perspect.* 124, 321–328. 10.1289/ehp.1408164. [PubMed: 26161573]
- Juni P, Witschi A, Bloch R, Egger M, 1999. The hazards of scoring the quality of clinical trials for meta-analysis. *JAMA* 282, 1054–1060. [PubMed: 10493204]
- Kasper-Sonnenberg M, Wittsiepe J, Wald K, Koch HM, Wilhelm M, 2017. Prepubertal exposure with phthalates and bisphenol A and pubertal development. *PLoS ONE* 12, e0187922. 10.1371/journal.pone.0187922.
- Kay VR, Chambers C, Foster WG, 2013. Reproductive and developmental effects of phthalate diesters in females [review]. *Crit. Rev. Toxicol.* 43, 200–219. 10.3109/10408444.2013.766149. [PubMed: 23405971]
- Languon I, Saillenfait AM, Payan JP, 1998. Comparative embryotoxicity of di-n-butyl phthalate and its main metabolite mono-n-butyl phthalate at midgestation. *Teratology* 58, 22A.
- Latini G, De Felice C, Presta G, Del Vecchio A, Paris I, Ruggieri F, Mazzeo P, 2003. In utero exposure to di-(2-ethylhexyl)phthalate and duration of human pregnancy. *Environ. Health Perspect.* 111, 1783–1785. 10.1289/ehp.6202. [PubMed: 14594632]
- Lee SW, Lee WJ, Chae HW, Lee EB, Kim JH, Kim DH, Kim HS, 2009. Determination of Serum Di-(2-ethylhexyl) Phthalate and Bisphenol A Level in Children With Idiopathic Central Precocious Puberty. 14. pp. 154–162.
- Li B, Xu X, Zhu Y, Cao J, Zhang Y, Huo X, 2016. Neonatal phthalate ester exposure induced placental MTs, FATP1 and HFABP mRNA expression in two districts of southeast China. *Sci. Rep.* 6, 21004. 10.1038/srep21004. [PubMed: 26867681]
- Liao KW, Kuo PL, Huang HB, Chang JW, Chiang HC, Huang PC, 2018. Increased risk of phthalates exposure for recurrent pregnancy loss in reproductive-aged women. *Environ. Pollut.* 241, 969–977. 10.1016/j.envpol.2018.06.022. [PubMed: 30029331]

- Lomenick JP, Calafat AM, Melguizo Castro MS, Mier R, Stenger P, Foster MB, Wintergerst KA, 2010. Phthalate exposure and precocious puberty in females. *J. Pediatr.* 156, 221–225. 10.1016/j.jpeds.2009.09.047. [PubMed: 19892364]
- Lovekamp-Swan T, Davis BJ, 2003. Mechanisms of phthalate ester toxicity in the female reproductive system [review]. *Environ. Health Perspect.* 111, 139–145. [PubMed: 12573895]
- Machtinger R, Gaskins AJ, Racowsky C, et al. , 2018. Urinary concentrations of biomarkers of phthalates and phthalate alternatives and IVF outcomes. *Environ. Int.* 111, 23–31 (PMID:29161633). [PubMed: 29161633]
- Mariana M, Feiteiro J, Verde I, Cairrao E, 2016. The effects of phthalates in the cardiovascular and reproductive systems: a review [review]. *Environ. Int.* 94, 758–776. 10.1016/j.envint.2016.07.004. [PubMed: 27424259]
- Martin JA, Hamilton BE, Osterman MJ, Curtin SC, Matthews TJ, 2013. Births: final data for 2012. *Natl. Vital Stat. Rep.* 62, 1–68.
- Meeker JD, Hu H, Cantonwine DE, Lamadrid-Figueroa H, Calafat AM, Ettinger AS, Hernandez-Avila M, Loch-Carusio R, Tellez-Rojo MM, 2009. Urinary phthalate metabolites in relation to preterm birth in Mexico city. *Environ. Health Perspect.* 117, 1587–1592. 10.1289/ehp.0800522. [PubMed: 20019910]
- Messerlian C, Souter I, Gaskins AJ, Williams PL, Ford JB, Chiu YH, Calafat AM, Hauser R, Team ES, 2015. Urinary phthalate metabolites and ovarian reserve among women seeking infertility care. *Hum. Reprod.* 10.1093/humrep/.
- Messerlian C, Wylie BJ, Minguez-Alarcon L, Williams PL, Ford JB, Souter IC, Calafat AM, Hauser R, Team ES, 2016. Urinary concentrations of phthalate metabolites and pregnancy loss among women conceiving with medically assisted reproduction. *Epidemiology* 27, 879–888. 10.1097/EDE.0000000000000525. [PubMed: 27299194]
- Mouritsen A, Frederiksen H, Sørensen K, Aksglaede L, Hagen C, Skakkebaek N, Main K, Andersson A, Juul A, 2013. Urinary phthalates from 168 girls and boys measured twice a year during a 5 year period: associations with adrenal androgen levels and puberty. *J. Clin. Endocrinol. Metab.* 98, 3755–3764. 10.1210/jc.2013-1284. [PubMed: 23824423]
- Mu D, Gao F, Fan Z, Shen H, Peng H, Hu J, 2015. Levels of phthalate metabolites in urine of pregnant women and risk of clinical pregnancy loss. *Environ. Sci. Technol.* 49, 10651–10657. 10.1021/acs.est.5b02617. [PubMed: 26251123]
- NAS, 2017. Application of Systematic Review Methods in an Overall Strategy for Evaluating Low-Dose Toxicity from Endocrine Active Chemicals. 10.17226/24758. <https://www.nap.edu/catalog/24758/application-of-systematicreview-methods-in-an-overall-strategy-for-evaluating-low-dose-toxicity-from-endocrine-active-chemicals>.
- Philips EM, Kahn LG, Jaddoe VWV, Shao Y, Asimakopoulos AG, Kannan K, Steegers EAP, Trasande L, 2018. First trimester urinary bisphenol and phthalate concentrations and time to pregnancy: a population-based cohort analysis. *J. Clin. Endocrinol. Metab.* 103, 35403547. 10.1210/jc.2018-00855.
- Pola ska K, Ligocka D, Sobala W, Hanke W, 2016. Effect of environmental phthalate exposure on pregnancy duration and birth outcomes. *Int. J. Occup. Med. Environ. Health* 29, 683–697. 10.13075/ijomeh.1896.00691. [PubMed: 27443763]
- Rais-Bahrami K, Nunez S, Revenis ME, Luban NL, Short BL, 2004. Follow-up study of adolescents exposed to di(2-ethylhexyl) phthalate (DEHP) as neonates on extracorporeal membrane oxygenation (ECMO) support. *Environ. Health Perspect.* 112, 1339–1340. 10.1289/ehp.6901. [PubMed: 15345350]
- Shi H, Cao Y, Shen Q, Zhao Y, Zhang Z, Zhang Y, 2015. Association between urinary phthalates and pubertal timing in Chinese adolescents. *J. Epidemiol.* 25, 574–582. 10.2188/jea.JE20140205. [PubMed: 26212725]
- Shoaff JR, Romano ME, Yolton K, Lanphear BP, Calafat AM, Braun JM, 2016. Prenatal phthalate exposure and infant size at birth and gestational duration. *Environ. Res.* 150, 52–58. 10.1016/j.envres.2016.05.033. [PubMed: 27236572]

- Smarr MM, Grantz KL, Sundaram R, Maisog JM, Kannan K, Louis GM, 2015. Parental urinary biomarkers of preconception exposure to bisphenol A and phthalates in relation to birth outcomes. *Environ. Health* 14, 73. 10.1186/s12940-015-0060-5. [PubMed: 26362861]
- Smerieri A, Testa C, Lazzeroni P, Nuti F, Grossi E, Cesari S, Montanini L, Latini G, Bernasconi S, Papini AM, Street ME, 2015. Di-(2-ethylhexyl) phthalate metabolites in urine show age-related changes and associations with adiposity and parameters of insulin sensitivity in childhood. *PLoS ONE* 10, e0117831. 10.1371/journal.pone.0117831.
- Specht IO, Bonde JP, Toft G, Lindh CH, Jönsson BA, Jørgensen KT, 2015. Serum phthalate levels and time to pregnancy in couples from Greenland, Poland and Ukraine. *PLoS ONE* 10, e0120070. 10.1371/journal.pone.0120070.
- Srilanchakon K, Thadsri T, Jantarat C, Thengyai S, Nosoognoen W, Supornsilchai V, 2017. Higher phthalate concentrations are associated with precocious puberty in normal weight Thai girls. *J. Pediatr. Endocrinol. Metab.* 30, 1293–1298. 10.1515/jpem-2017-0281. [PubMed: 29176028]
- Sterne J, Higgins J, Reeves B, 2016. ROBINS-I: a tool for assessing risk of bias in nonrandomized studies of interventions, version 7 March 2016. *Br. Med. J.* 355, i4919. [PubMed: 27733354]
- Su PH, Chen JY, Lin CY, Chen HY, Liao PC, Ying TH, Wang SL, 2014. Sex steroid hormone levels and reproductive development of eight-year-old children following in utero and environmental exposure to phthalates. *PLoS ONE* 9, e102788. 10.1371/journal.pone.0102788.
- Suzuki Y, Niwa M, Yoshinaga J, Mizumoto Y, Serizawa S, Shiraishi H, 2010. Prenatal exposure to phthalate esters and PAHs and birth outcomes. *Environ. Int.* 36, 699–704. 10.1016/j.envint.2010.05.003. [PubMed: 20605637]
- Thomsen AM, Riis AH, Olsen J, Jönsson BA, Lindh CH, Hjollund NH, Jensen TK, Bonde JP, Toft G, 2017. Female exposure to phthalates and time to pregnancy: a first pregnancy planner study. *Hum. Reprod.* 32, 232–238. 10.1093/humrep/dew291. [PubMed: 27852689]
- Toft G, Jönsson BA, Lindh CH, Jensen TK, Hjollund NH, Vested A, Bonde JP, 2012. Association between pregnancy loss and urinary phthalate levels around the time of conception. *Environ. Health Perspect.* 120, 458–463. 10.1289/ehp.1103552. [PubMed: 22113848]
- U.S. EPA (U.S. Environmental Protection Agency), 1996. Guidelines for Reproductive Toxicity Risk Assessment. Risk Assessment Forum U.S. Environmental Protection Agency, Washington, DC, pp. 1–143(EPA/630/R-96/009). (https://www.epa.gov/sites/production/files/2014-11/documents/guidelines_repro_toxicity.pdf).
- Vandenberg LN, Colborn T, Hayes TB, Heindel JJ, Jacobs DR, Lee DH, Shioda T, Soto AM, vom Saal FS, Welshons WV, Zoeller RT, Myers JP, 2012. Hormones and endocrine-disrupting chemicals: low-dose effects and nonmonotonic dose responses [review]. *Endocr. Rev.* 33, 378–455. 10.1210/er.2011-1050. [PubMed: 22419778]
- Vélez MP, Arbuckle TE, Fraser WD, 2015. Female exposure to phenols and phthalates and time to pregnancy: the Maternal-Infant Research on Environmental Chemicals (MIREC) study. *Fertil. Steril.* 103, 1011–1020.e1012. 10.1016/j.fertnstert.2015.01.005. [PubMed: 25681860]
- Watkins DJ, Téllez-Rojo MM, Ferguson KK, Lee JM, Solano-Gonzalez M, Blank-Goldenberg C, Peterson KE, Meeker JD, 2014. In utero and peripubertal exposure to phthalates and BPA in relation to female sexual maturation. *Environ. Res.* 134, 233–241. 10.1016/j.envres.2014.08.010. [PubMed: 25173057]
- Watkins DJ, Milewski S, Domino SE, Meeker JD, Padmanabhan V, 2016. Maternal phthalate exposure during early pregnancy and at delivery in relation to gestational age and size at birth: a preliminary analysis. *Reprod. Toxicol.* 65, 59–66. 10.1016/j.reprotox.2016.06.021. [PubMed: 27352641]
- Watkins DJ, Sánchez BN, Téllez-Rojo MM, Lee JM, Mercado-García A, Blank-Goldenberg C, Peterson KE, Meeker JD, 2017. Phthalate and bisphenol A exposure during in utero windows of susceptibility in relation to reproductive hormones and pubertal development in girls. *Environ. Res.* 159, 143–151. 10.1016/j.envres.2017.07.051. [PubMed: 28800472]
- Weinberger B, Vetrano AM, Archer FE, Marcella SW, Buckley B, Wartenberg D, Robson MG, Klim J, Azhar S, Cavin S, Wang L, Rich DQ, 2014. Effects of maternal exposure to phthalates and bisphenol A during pregnancy on gestational age. *J. Matern. Fetal Neonatal Med.* 27, 323–327. 10.3109/14767058.2013.815718. [PubMed: 23795657]

- Whyatt RM, Adibi JJ, Calafat AM, Camann DE, Rauh V, Bhat HK, Perera FP, Andrews H, Just AC, Hoepner L, Tang D, Hauser R, 2009. Prenatal di(2-ethylhexyl)phthalate exposure and length of gestation among an inner-city cohort. *Pediatrics* 124, e1213–e1220. 10.1542/peds.2009-0325. [PubMed: 19948620]
- Wolff MS, Engel SM, Berkowitz GS, Ye X, Silva MJ, Zhu C, Wetmur J, Calafat AM, 2008. Prenatal phenol and phthalate exposures and birth outcomes. *Environ. Health Perspect.* 116, 1092–1097. 10.1289/ehp.11007. [PubMed: 18709157]
- Wolff MS, Teitelbaum SL, Pinney SM, Windham G, Liao L, Biro F, Kushi LH, Erdmann C, Hiatt RA, Rybak ME, Calafat AM, 2010. Investigation of relationships between urinary biomarkers of phytoestrogens, phthalates, and phenols and pubertal stages in girls. *Environ. Health Perspect.* 118, 1039–1046. 10.1289/ehp.0901690. [PubMed: 20308033]
- Wolff MS, Teitelbaum SL, McGovern K, Windham GC, Pinney SM, Galvez M, Calafat AM, Kushi LH, Biro FM, 2014. Phthalate exposure and pubertal development in a longitudinal study of US girls. *Hum. Reprod.* 29, 1558–1566. 10.1093/humrep/deu081. [PubMed: 24781428]
- Wolff MS, Pajak A, Pinney SM, Windham GC, Galvez M, Rybak M, Silva MJ, Ye X, Calafat AM, Kushi LH, Biro FM, Teitelbaum SL, Program, BCaER, 2017. Associations of urinary phthalate and phenol biomarkers with menarche in a multiethnic cohort of young girls. *Reprod. Toxicol.* 67, 56–64. 10.1016/j.reprotox.2016.11.009. [PubMed: 27851993]
- Wu H, Ashcraft L, Whitcomb BW, Rahil T, Tougias E, Sites CK, Pilsner JR., 2017. Parental contributions to early embryo development: influences of urinary phthalate and phthalate alternatives among couples undergoing IVF treatment. *Hum. Reprod.* 32, 6.
- Yi H, Gu H, Zhou T, Chen Y, Wang G, Jin Y, Yuan W, Zhao H, Zhang L, 2016. A pilot study on association between phthalate exposure and missed miscarriage. *Eur. Rev. Med. Pharmacol. Sci.* 20, 1894–1902. [PubMed: 27212185]
- Yum T, Lee S, Kim Y, 2013. Association between precocious puberty and some endocrine disruptors in human plasma. *J. Environ. Sci. Health A Tox. Hazard. Subst. Environ. Eng.* 48, 912–917. 10.1080/10934529.2013.762734. [PubMed: 23485241]
- Zhang Y, Cao Y, Shi H, Jiang X, Zhao Y, Fang X, Xie C, 2015. Could exposure to phthalates speed up or delay pubertal onset and development? A 1.5-year follow-up of a school-based population. *Environ. Int.* 83, 41–49. 10.1016/j.envint.2015.06.005. [PubMed: 26073845]
- Zhao R, Wu Y, Zhao F, Lv Y, Huang D, Wei J, Ruan C, Huang M, Deng J, Huang D, Qiu X, 2017. The risk of missed abortion associated with the levels of tobacco, heavy metals and phthalate in hair of pregnant woman: a case control study in Chinese women. *Medicine (Baltimore)* 96, e9388. 10.1097/MD.0000000000009388. [PubMed: 29390543]

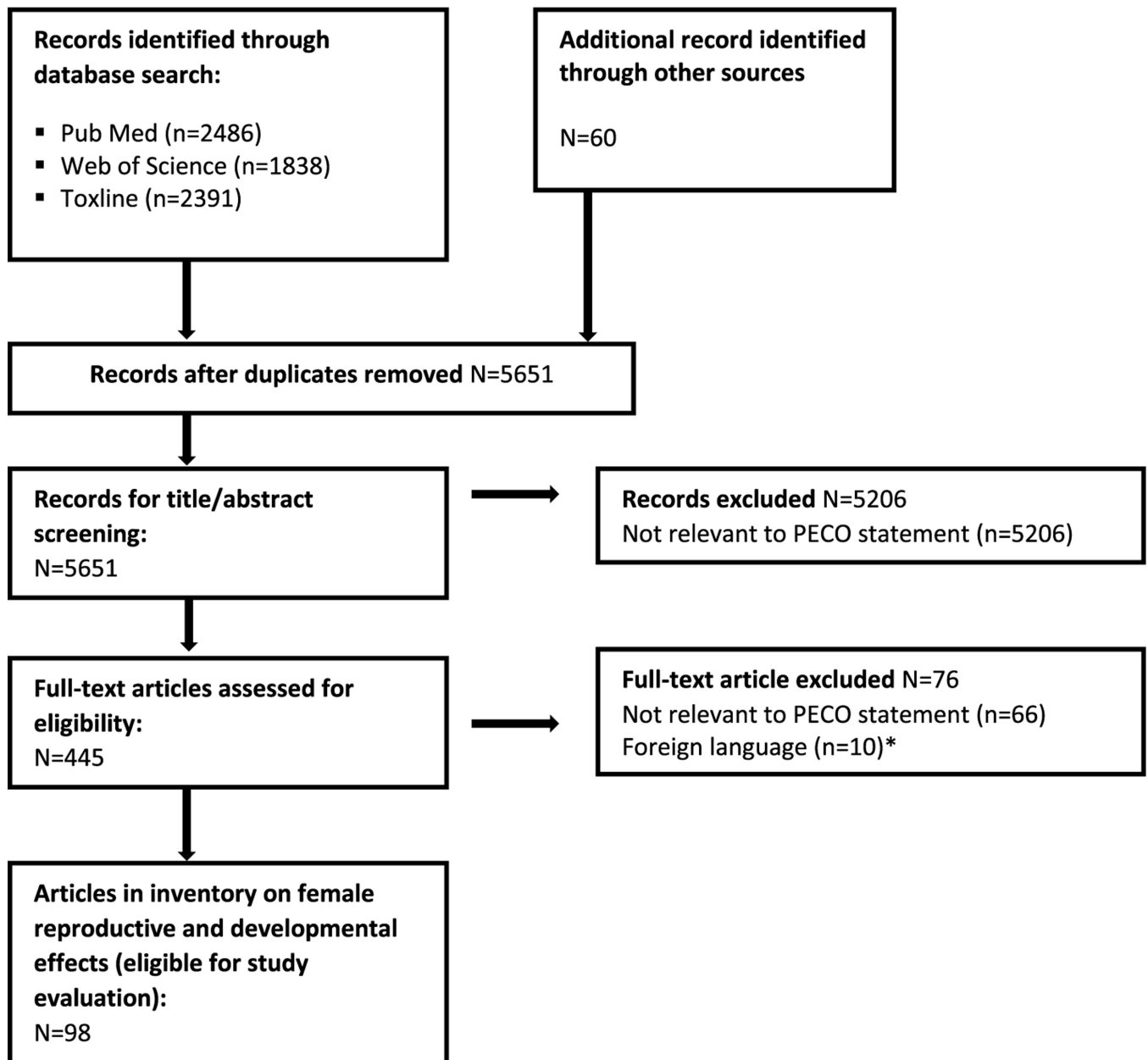
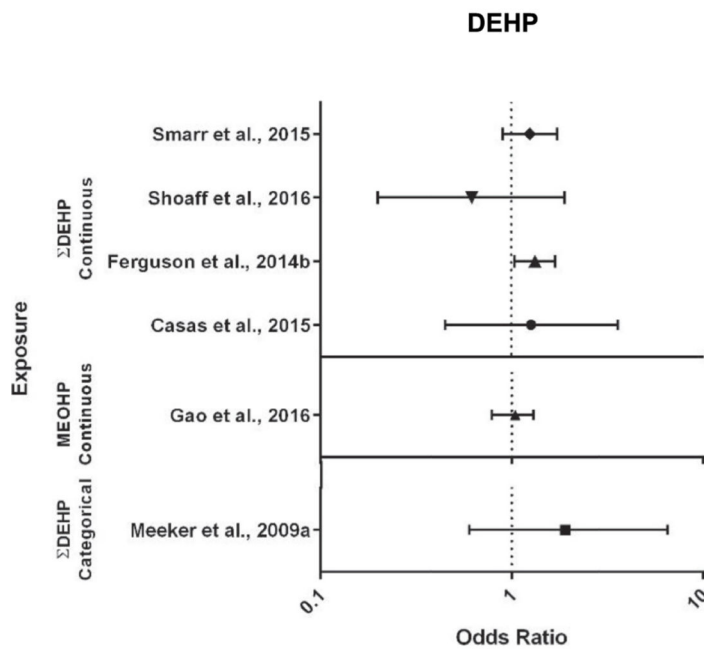
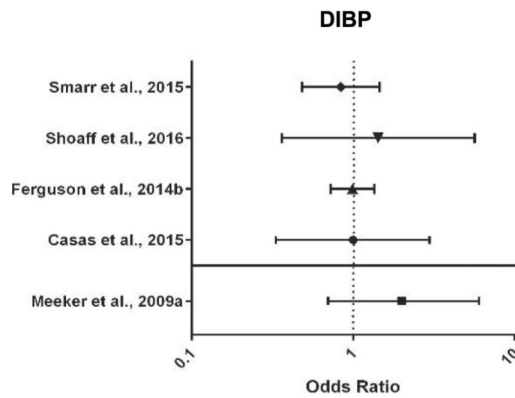
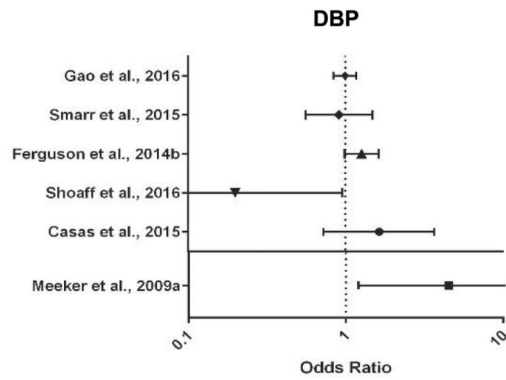


Fig. 1. Literature flow diagram for female reproductive and developmental effects of phthalates.
*Included one study on low birth weight and one on precocious puberty. Based on the abstracts, both would be excluded due to temporality issues.

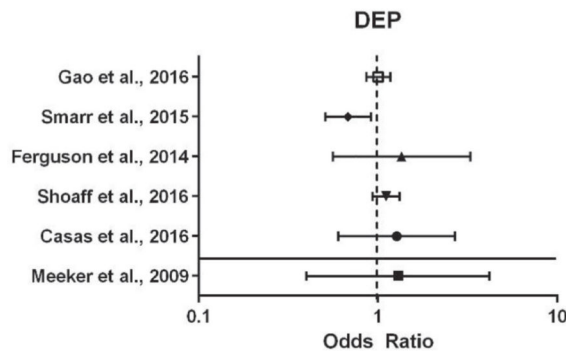
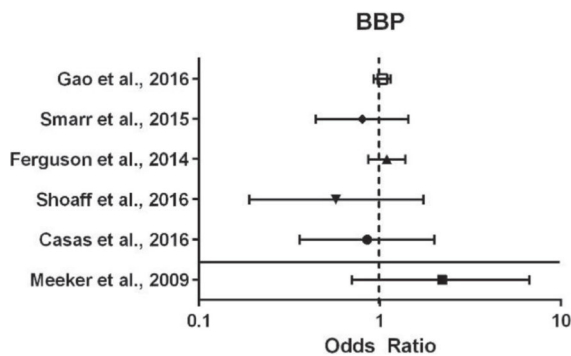


Reference; Study Confidence; N	Mean (IQR) (nm/mL)	OR (95% CI) (continuous In-transformed exposure with 1 unit change or as specified)
ΣDEHP		
Smarr et al. (2015); Medium; 200	0.2 (75 th : 0.3)	1.25 (0.90, 1.74) Quartiles Q2: 0.77 (0.26, 2.28) Q3: 1.38 (0.46, 4.13) Q4: 0.93 (0.29, 2.97)
Shoaff et al. (2016); High; 368	0.3 (0.1-0.7)	Continuous, 10-fold change 0.62 (0.20, 1.90)
Ferguson et al. (2014b); High; 482	0.4	1.33 (1.04, 1.70) Quartiles Q2: 1.23 Q3: 1.55 Q4: 2.17 <i>p</i> trend = 0.02*
Casas et al. (2016); High; 380	93 ng/mL (range 14-1,429)	1.27 (0.45, 3.62) Quartiles Q2: 1.80 (0.15, 21.33) Q3: 3.32 (0.33, 33.81) Q4: 2.45 (0.23, 25.79)
Meeker et al. (2009b); Medium; 60	0.2 (0.1-0.4)	Dichotomous 1.9 (0.6, 6.5)
MEOHP		
Gao et al. (2016); Medium; 3103	7.5 (4.7-29) ng/mL	1.02 (0.80, 1.31)

Fig. 2. Association between DEHP exposure and preterm birth.
 **p* < 0.05, results that support an association are shaded. Dark gray represents one or more of the following: *p* < 0.05, large effect size (e.g., OR 1.5), or exposure-response trend across categories of exposure. Light gray represents other supportive results.



Reference; Study Confidence; N	Exposure mean (IQR) (ng/mL)	OR (95% CI) (continuous ln-transformed exposure with 1 unit change or as specified)
DBP (measured by MBP)		
Smarr et al. (2015); Medium; 2000	6.8 (75 th : 15.8)	0.91 (0.56, 1.49) Quartiles Q2: 0.66 (0.23, 1.87) Q3: 0.74 (0.21, 2.57) Q4: 0.68 (0.17, 2.65)
Ferguson et al. (2014b); High; 368	5.7 (2.2-11)	1.27 (0.99, 1.63) Quartiles Q2: 0.82 Q3: 1.57 Q4: 2.25 <i>p</i> trend = 0.03
Shoaff et al. (2016); High; 368	27 (23-54)	Continuous, 10-fold change 0.20 (0.04, 0.96)
Casas et al. (2016); High; 380	29 (3.4-403)	1.64 (0.73, 3.67)
Meeker et al. (2009); Medium; 60	33 (21-74)	Dichotomous 4.5 (1.2, 16.6)
Gao et al. (2016); Medium; 3103	42 (24-319)	0.995 (0.84, 1.18)
DIBP (measured by MiBP)		
Smarr et al. (2015); Medium; 200	3.7 (75 th : 7.7)	0.83 (0.48, 1.44) Quartiles Q2: 1.31 (0.26, 6.65) Q3: 1.21 (0.24, 6.14) Q4: 0.72 (0.06, 9.18)
Shoaff et al. (2016); High; 368	5.7 (2.2-11)	Continuous, 10-fold change 1.41 (0.36, 5.58)
Ferguson et al. (2014b); High; 482	6.8 (4.5-10)	0.98 (0.72, 1.34)
Casas et al. (2016); High; 380	29 (range 4.0-368)	0.99 (0.33, 2.94) Quartiles Q2: 0.94 (0.12, 7.18) Q3: 1.97 (0.33, 11.8) Q4: 0.42 (0.04, 4.99)
Meeker et al. (2009b); Medium; 60	1.9 (0.8-4.1)	Dichotomous 2.0 (0.7, 6.0)



BBP (measured by MBzP)		
Gao et al. (2016); Medium; 3103	0.09 (<LOD-1.1)	1.03 (0.92,1.14)
Smarr et al. (2015); Medium; 200	3.5 (75 th : 8.5)	0.80 (0.44, 1.44) Quartiles Q2: 0.43 (0.14, 1.27) Q3: 0.85 (0.27, 2.71) Q4: 0.51 (0.12, 2.17)
Ferguson et al. (2014b); High; 482	6.5 (3.3-12)	1.09 (0.86, 1.38)
Shoaff et al. (2016); High; 268	10 (4.2-24)	Continuous, 10-fold change 0.57 (0.19, 1.74)
Casas et al. (2016); High; 380	11 (range 0.8-336)	0.85 (0.36, 2.00) Quartiles Q2: 0.75 (0.12, 4.87) Q3: 0.70 (0.11, 4.42) Q4: 0.57 (0.09, 3.70)
Meeker et al. (2009b); Medium; 60	2.3 (1.0-5.2)	Dichotomous 2.2 (0.7, 6.7)
DEP (measured by MEP)		
Gao et al. (2016); Medium; 3013	8.8 (4.3-75)	1.002 (0.86,1.17)
Smarr et al. (2015); Medium; 200	101 (75 th : 194)	0.68 (0.51, 0.91) Quartiles Q2: 0.44 (0.15, 1.25) Q3: 0.72 (0.25, 2.04) Q4: 0.26 (0.07, 0.89)
Ferguson et al. (2014b); High; 368	131 (54-357)	Continuous, 10-fold change 1.35 (0.56, 3.27)
Shoaff et al. (2016); High; 482	134	1.11 (0.93, 1.32)
Casas et al. (2016); High; 380	336 (range 22-5,115)	1.27 (0.60, 2.68)
Meeker et al. (2009b); Medium; 60	108	Dichotomous 1.3 (0.4, 4.2)

Fig. 3. Association between exposure to DBP, DIBP, BBP, and DEP and preterm birth. Each panel depicts results for a metabolite(s) from a different parent phthalate; Q= quartile. * $p < 0.05$, results that support an association are shaded. Dark gray represents one or more of the following: $p < 0.05$, large effect size (e.g., OR 1.5), or exposure-response trend across categories of exposure. Light gray represents other supportive results. Tables are sorted within individual phthalates by median exposure level.

Table 1

Outcomes included in the systematic review.

Outcome	Background and relevance to female reproductive and/or developmental toxicity
Pubertal development	<ul style="list-style-type: none"> • Puberty is a continuous process involving maturation of both the hypothalamic-pituitary-adrenal axis and the hypothalamic-pituitary-gonadal axis. • Hormonal changes during puberty underlie the natural dynamic period of physical and sexual maturation that culminates in the ability to reproduce. • Changes in the mean age at pubertal onset in either direction would be considered an adverse effect (U.S. EPA, 1996).
Fecundity	<ul style="list-style-type: none"> • Fecundity is the biological capacity to reproduce. This review includes time to pregnancy as the primary outcome measure used to study fecundity, as well as rate of clinical pregnancy. • Time to pregnancy displays a reverse J-shaped cumulative distribution with many couples (30–40%) conceiving in the first cycle (Buck Louis et al., 2011) and 80–90% of couples being pregnant after up to 12 months of trying (Gnoth et al., 2003). • Fecundity within a couple can be influenced by either male or female exposures, or both (this review focuses on female exposure). Increases in time to pregnancy are considered indicative of reproductive toxicity (U.S. EPA, 1996). • Related outcomes among women seeking infertility treatment are summarized for coherence (without systematic review): oocytes, follicle count, embryo quality, and implantation.
Spontaneous abortion	<ul style="list-style-type: none"> • Spontaneous abortion or miscarriage is a pregnancy loss occurring before 20 weeks of gestation. • Loss that occurs before the pregnancy is clinically recognized is considered early, or preclinical. • Losses at 5 to roughly 20 weeks gestation are considered clinical loss.
Preterm birth	<ul style="list-style-type: none"> • Preterm birth, a delivery occurring before 37 weeks gestation, is a relatively common complication of pregnancy, occurring in > 10% of births (Martin et al., 2013). • Changes in gestational duration are considered indicative of reproductive toxicity (U.S. EPA, 1996) and are also reviewed.

Table 2

Epidemiology studies of pubertal development in females.

Reference	Study description				Includes metabolites of:						Study evaluation					
	Population	Exposure	Outcome	DEH- P	DIN- P	DB- P	DIB- P	BB- P	DE- P	Exposure	Outcome	Selection	Confounding	Analysis	Overall confidence	
Included Hart et al. (2013)	Pregnancy cohort (N = 123 girls) in Australia	Blood samples at 18 and 34/36 wks gestation	Prospectively reported age at menarche	✓	✓	✓ ^a	✓ ^a	✓ ^a	✓ ^a	D	G	D	D	D	Low	
Mouritsen et al. (2013)	Cohort of children (N=84 girls) in Denmark enrolled at 6–12 yrs	Urine samples every 6 mo, at least 2 samples averaged	Tanner stages determined by pediatrician	✓	✓	✓ ^b	✓ ^b	✓	✓	G	G	D	D	D	Low	
Su et al. (2014)	Pregnancy cohort (N = 69 girls) in Taiwan	Single urine sample (3 rd trimester)	Tanner stages determined by single examiner	✓	✓	✓	✓	✓	✓	A/D	G	D	D	D	Low	
Watkins et al. (2017); Watkins et al. (2014)	Pregnancy cohort (N = 117 girls) in Mexico	Urine sample in each trimester	Tanner stages determined by pediatrician	✓	✓	✓	✓	✓	✓	G	G	D	A	A	Medium	
Wolff et al. (2017); Wolff et al. (2014); Wolff et al. (2010)	Cohort of girls (N= 1,170) in U.S. enrolled at 6–8 yrs	Single urine sample at baseline	Repeated assessments of Tanner stages by trained examiners	✓	✓	✓	✓	✓	✓	A/D	G	A	G	A	Medium	
Total Studies per Phthalate				7	3	6	5	6	6							

Excluded studies: Exposure measured after development of outcome: Buttke et al. (2012), Chen et al. (2013), Chou et al. (2009), Frederiksen et al. (2012), Hashemipour et al. (2018), Lomenick et al. (2010), Shi et al. (2015), Srilanchakon et al. (2017), Zhang et al. (2015). Exposure measured in blood: Bulu et al. (2016), Lee et al. (2009), Yum et al. (2013). Exposure measured by self-report: Garrao and Shaw (2012). No comparison group: Rais-Bahrani et al. (2004). Critical deficiency in population selection: Smerterti et al. (2015).

Additional studies identified in the August 2018 literature update: Binder et al. (2018), Kasper-Sonnenberg et al. (2017), Watkins et al. (2017), an update of Watkins et al. (2014) was also identified in the search and incorporated into the synthesis.

G = good; A = adequate; D = deficient; A/P = adequate for short chain phthalates, deficient for long chain phthalates.

^aStudy was considered *critically deficient* for this phthalate due to measurement in blood.

^bStudy summed DBP and DIBP metabolites.

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript

Table 3

Associations between DEHP exposure and pubertal development.

Reference; Study confidence; N	Outcome	Effect estimate ^a	Mean exposure (IQR) (ng/mL)	Pubertal development timing effect estimate
DEHP				
Childhood Exposure				
Wolff et al. (2017) and Wolff et al. (2014); <u>Medium</u> : 1,141	Age at menarche (n=1,051) Age at 1 st breast stage 2 Age at 1 st pubic hair stage 2	HR (95% CI)	Median 142 µg/g creatinine	1.00 (0.94, 1.06) 1.01 (0.94, 1.09) 0.92 (0.85, 0.99) *
Mouritsen et al. (2013); <u>Low</u> : 84	Age at breast stage 2+ Age at pubic hair stage 2+	Medians in high vs. low exposure	115 (14–4,627)	“no significant difference”
Prenatal Exposure				
Watkins et al. (2014); <u>Medium</u> : 120	Menarche Breast stage 2+ Pubic hair stage 2+	OR (95% CI)	0.24 µmol/L	1.89 (0.57, 6.22) 0.66 (0.26, 1.69) 2.15 (0.62, 7.48)
Hart et al. (2013); <u>Low</u> : 121; <u>Su et al. (2014)</u> ; <u>Low</u> : 69	Age at menarche Tanner stage > 1	Correlation coefficient OR (95% CI)	4.3 (2.7–7.1) 47 mg/g creatinine	–0.17, <i>p</i> = 0.07 Stage 2: 1.00 (1.00, 1.00) Stage 3: 1.00 (0.99, 1.01)

^a OR and hazard ratio (HR) < 1 indicate later pubertal development. Bolding of the effect estimate indicates evidence that suggests later onset of puberty, while grey shading indicates evidence that suggests earlier onset of puberty. NR = not reported; Q = quartile, y/n = yes/no

* *p* < 0.05.

Table 4

Associations between DBP and DIBP exposure and pubertal development.

Reference; Study Confidence; N	Outcome	Effect estimate	Mean exposure (IQR) (ng/mL)	Pubertal development timing effect estimate
DBP				
Childhood Exposure				
Wolff et al. (2017) and Wolff et al. (2014); <u>Low</u> : 1,141	Age at menarche (n=1,051) Age at 1 st breast stage 2 Age at 1 st pubic hair stage 2	HR (95% CI)	Median 14 µg/g creatinine	0.98 (0.90, 1.06) 0.99 (0.92, 1.08) 0.92 (0.85, 1.00)
Prenatal Exposure				
Su et al. (2014); <u>Low</u> : 69	Tanner stage > 1	OR (95% CI)	68 mg/g creatinine	Stage 2: 1.00 (0.99, 1.00) Stage 3: 0.99 (0.98, 1.01)
Watkins et al. (2014); <u>Medium</u> : 116	Menarche Breast stage 2+ Pubic hair stage 2+	OR (95% CI)	57 (26–119) in Watkins et al., 2014	0.73 (0.22,2.41) 1.06 (0.41,2.71) 2.6 (0.65,10.46)
DBP + DIBP				
Childhood Exposure				
Mouritsen et al. (2013); <u>Low</u> : 84	Age at breast stage 2+ Age at pubic hair stage 2+	Medians in high vs. low exposure	122 (range 23–904)	Not significant, slightly lower age at pubarche Not significant, slightly higher age at pubarche
DIBP				
Childhood Exposure				
Wolff et al. (2017) and Wolff et al. (2014); <u>Medium</u> : 1,141	Age at menarche (n=1,051) Age at 1 st breast stage 2 Age at 1 st pubic hair stage 2	HR (95% CI)	Median 48 µg/g creatinine	1.01 (0.94, 1.09) 0.97 (0.9, 1.04) 0.99 (0.91, 1.07)
Prenatal Exposure				
Watkins et al. (2014); <u>Low</u> : 116	Menarche Breast stage 2+ Pubic hair stage 2+	OR (95% CI)	2 (1.0–3.6) in Watkins et al., 2014	0.39 (0.09,1.70) 1.21 (0.38,3.90) 1.65 (0.30,9.03)

^aOR and hazard ratio (HR) < 1 indicate later pubertal development

^bIncludes MBP and OH-MBP

^cIncludes MIBP and OH-MIBP, Bolding of the effect estimate indicates evidence that suggests later onset of puberty, while grey shading indicates evidence that suggests earlier onset of puberty. NR = not reported; Q = quartile

* p < 0.05

Table 5

Associations between BBP and DEP exposure and pubertal development.

Reference; Study Confidence; N	Outcome	Effect estimate	Mean exposure (IQR) (ng/mL)	Pubertal development timing effect estimate
BBP				
Childhood Exposure				
Wolff et al. (2017) and Wolff et al. (2014); <u>Medium</u> : 1,141	Age at menarche (n=1,051) Age at 1 st breast stage 2 Age at 1 st pubic hair stage 2	HR (95% CI)	Median 21 µg/g creatinine	0.98 (0.92, 1.05) 0.95 (0.89, 1.01) 0.92 (0.86, 0.99) *
Mouritsen et al. (2013); <u>Low</u> : 84	Age at breast stage 2+ Age at pubic hair stage 2+	Medians in high vs. low exposure	37 (range 3–433)	Not significant, slightly lower age at pubarche “no significant difference”
Prenatal Exposure				
Su et al. (2014); <u>Low</u> : 69	Tanner stage > 1	OR (95% CI)	17 mg/g	Stage 2: 1.01 (0.99, 1.02) Stage 3: 0.98 (0.91, 1.05)
Watkins et al. (2014); <u>Low</u> : 116	Menarche Breast stage 2+ Pubic hair stage 2+	OR (95% CI)	4 (2.2–7.0) in Watkins et al., 2014	0.92 (0.26,3.25) 1.58 (0.63,3.96) 3.89 (0.96,15.8)
DEP				
Childhood Exposure				
Wolff et al. (2017) and Wolff et al. (2014); <u>Medium</u> : 1,141	Age at menarche (n=1,051) Age at 1 st breast stage 2 Age at 1 st pubic hair stage 2	HR (95% CI)	Median 68 µg/g creatinine	1.06 (0.99, 1.14) 1.03 (0.96, 1.09) 0.99 (0.94, 1.06)
Mouritsen et al. (2013); <u>Low</u> : 84	Age at breast stage 2+ Age at pubic hair stage 2+	Medians in high vs. low exposure	36 (range 6–847)	“no significant difference”
Prenatal Exposure				
Su et al. (2014); <u>Low</u> : 69	Tanner stage > 1	OR (95% CI)	65 mg/g creatinine	Stage 2: 1.00 (1.00, 1.01) Stage 3: 0.99 (0.94, 1.03)
Watkins et al. (2014); <u>Low</u> : 116	Menarche Breast stage 2+ Pubic hair stage 2+	OR (95% CI)	114 (41–242) in Watkins et al., 2014	4.33 (1.25,15.0) 1.49 (0.57,3.88) 2.31 (0.67,7.94)

^aOR and hazard ratio (HR) < 1 indicate later pubertal development. Bolding of the effect estimate indicates evidence that suggests later onset of puberty, while grey shading indicates evidence that suggests earlier onset of puberty. Studies are sorted by median exposure level, NR = not reported; Q = quartile.

* p < 0.05

Table 6

Epidemiology studies of fecundity in females

Reference	Study description				Includes metabolites of:								Study evaluation				Overall confidence
	Population	Exposure	Outcome		DEH- P	DIN- P	DB- P	DIB- P	BB- P	DE- P	Exposure	Outcome	Selection	Confounding	Analysis		
Included Buck Louis et al. (2014)	Cohort (N = 501) in U.S. of couples trying to conceive	Single urine sample at study entry	Cycles to conception (measured prospectively)		✓	✓ ^a	✓	✓	✓	✓	A/D	G	G	A	A	Medium	
Jukic et al. (2016)	Cohort (N = 221) in U.S. of couples trying to conceive	Three urine samples per cycle, pooled	Cycles to conception (measured prospectively), up to 6 mo		✓	✓	✓	✓	✓	✓	G	G	A	G	A	High	
Thomsen et al. (2017)	Cohort (N = 229) in Denmark of couples trying to conceive	1–2 urine samples from final two cycles in study	Cycles to conception (measured prospectively), up to 6 mo		✓	✓	✓	✓	✓	✓	G/A	G	D	A	A	Medium	
Hauser et al. (2016)	Cohort (N = 256) in U.S. of women undergoing IVF	Two urine samples during each IVF (could contribute multiple cycles)	Rate of clinical pregnancy confirmed by ultrasound at ~6 wk		✓	✓	✓	✓	✓	✓	G	A	G	G	G	High	
Total Studies per Phthalate					4	2	4	3	4	4							

Excluded studies: Exposure measured after development of outcome: Specht et al. (2015), Vélez et al. (2015), Huang et al. (2016). Exposure categorized as occupational exposure to any phthalate (job exposure matrix): Burdorf et al. (2011). Additional studies identified in the August 2018 literature update: Phillips et al. (2018), Buck Louis et al. (2018), G=good; A=adequate; D=poor; A/D=adequate for short chain phthalates, deficient for long chain phthalates; IVF=in vitro fertilization.

^aStudy was considered *critically deficient* for this phthalate due to a high percent below the LOD.

Table 7

Associations between DEHP exposure and time to pregnancy.

Reference; study confidence; N	Mean exposure (IQR) (ng/mL)	Fecundability OR (95% CI)	Fecundability OR unit change
DEHP			
MEHP			
Buck Louis et al. (2014); Medium; 501	4.6	0.99 (0.87, 1.12)	1 SD change (log transformed)
Jukic et al. (2016); High; 221	6.6 (3.8–11)	T2: 1.23 (0.85, 1.79) T3: 1.39 (0.97, 1.97)	T2 and T3 vs. T1
Thomsen et al. (2017); Medium; 229	14.5 (range 0–348)	0.99 (0.72, 1.35)	1 unit increase (ln transformed)
MEOHP			
Buck Louis et al. (2014); Medium; 501	Mean 8.7	1.66 (0.91, 1.24)	1 SD change (log transformed)
Jukic et al. (2016); High; 221	33.3 (20–49)	Q2: 1.28 (0.91, 1.81) Q3: 1.11 (0.77, 1.58)	T2 and T3 vs. T1

Fecundability OR < 1 indicates increased time to pregnancy.

Q = quartile; T = tertile. Studies within individual phthalates are sorted by median exposure level.

Table 8

Associations between DINP, DBP, DIBP, BBP, and DEP exposure and time to pregnancy

Reference; Study confidence; N	Mean exposure (IQR) (ng/mL)	Fecundability OR (95% CI)	Fecundability OR unit change
DINP			
Jukic et al. (2016) ; High; 221	MCIOP 3.3 (2.3–5.2)	Q2: 0.93 (0.66, 1.32) Q3: 0.96 (0.68, 1.35)	T2 and T3 vs. T1
DBP			
Buck Louis et al. (2014); Medium; 501	10	0.93 (0.77, 1.12)	1 SD change (log transformed)
Jukic et al. (2016); High; 221	80 (49–127)	Q2: 0.90 (0.63, 1.27) Q3: 0.96 (0.68, 1.35)	T2 and T3 vs. T1
Thomsen et al. (2017); Medium; 229	272 (range 10–2,292)	1.03 (0.76, 1.39)	1 unit increase (ln transformed)
DIBP			
Jukic et al. (2016) ; High; 221	3.1 (1.9–5.0)	Q2: 1.16 (0.81, 1.65) Q3: 1.11 (0.77, 1.61)	Tertile vs. T1
Buck Louis et al. (2014) ; Medium; 501	5.1	0.95 (0.78, 1.14)	1 SD change (log transformed)
BBP			
Buck Louis et al. (2014) ; Medium; 501	5	0.94 (0.78, 1.13)	1 SD change (log transformed)
Thomsen et al. (2017) ; Medium; 229	22 (range 1.0–203)	0.88 (0.64, 1.19)	1 unit increase (ln transformed)
Jukic et al. (2016) ; High; 221	40 (25–69)	Q2: 0.76 (0.53, 1.10) Q3: 1.01 (0.73, 1.40)	T2 and T3 vs. T1
DEP			
Thomsen et al. (2017) ; Medium; 229	379 (range 16–12,357)	0.79 (0.63, 0.99) *	1 unit increase (ln transformed)
Buck Louis et al. (2014) ; Medium; 501	94	0.97 (0.84, 1.12)	1 SD change (log transformed)
Jukic et al. (2016) ; High; 221	134 (72–296)	Q2: 1.00 (0.71, 1.42) Q3: 0.98 (0.68, 1.40)	T2 and T3 vs. T1

Fecundability OR < 1 indicates increased time to pregnancy.

* p < 0.05, results that support an association are shaded. Q=quartile; T = tertile.

Table 9

Epidemiology studies of spontaneous abortion

Reference	Study description			Includes metabolites of:								Study evaluation				
	Population	Exposure	Outcome	DEH- P	DIN- P	DB- P	DIB- P	BB- P	DE- P	Exposure	Outcome	Selection	Confounding	Analysis	Overall confidence	
Included Jukic et al. (2016)	Preconception cohort in U.S. (N = 221 women)	Three urine samples from cycle of conception, pooled	Early pregnancy loss, identified via hCG	✓	✓	✓	✓	✓	✓	G	G	A	G	G	High	
Messertian et al. (2016)	Cohort of women receiving assisted reproductive technology in U.S. (N = 256)	Two urine samples per conception cycle	Total pregnancy loss and "biochemical" (early) loss identified prospectively	✓	✓	✓	✓	✓	✓	G	G	A	A	A	High	
Mu et al. (2015)	Case-control in China (N=132 cases, 172 controls) of women receiving ultrasound	Single morning urine sample at 5–13 wks gestation	Clinical pregnancy loss identified by ultrasound	✓	✓	✓	✓ ^a	✓	✓	A/P	A	D	D	A	Low	
Toft et al. (2012)	Preconception cohort in Denmark (N = 242 women)	Single urine sample from cycle of conception	Early and clinical pregnancy loss identified through urine samples or medical provider	✓	✓	✓	✓	✓	✓	A/P	G	D	A	A	Low	
Yi et al. (2016)	Case-control in China (N = 150 cases, 172 controls) of women	Single morning urine sample at admission to hospital	Clinical pregnancy low identified by ultrasound	✓	✓	✓	✓	✓	✓	A/P	A	D	A	A	Low	
Total Studies per Phthalate				5	2	5	4	3	5							

Additional studies identified in the August 2018 literature update: Liao et al. (2018), Zhao et al. (2017), Gao et al. (2017). G = good; A = adequate; D = deficient; A/D = adequate for short chain phthalates, deficient for long chain phthalates; hCG = human chorionic gonadotropin.

^aStudy was considered *critically deficient* for this phthalate due to a high percent below the LOD.

Table 10

Associations between DEHP metabolite levels and spontaneous abortion

Reference; Study Confidence; N	Early loss			Clinical loss		
	MEHP median (IQR), (ng/mL)	MEHP Early loss OR (95% CI)	MEOHHP Early loss OR (95% CI)	MEHP Clinical loss OR (95% CI)	MEOHHP Clinical loss OR (95% CI)	
Messerlian et al. (2016); High; 256	2.6 (1.5–6.4)	10.2	DEHP Biochemical loss Q2: 2.3 (0.63,8.5) Q3: 2.0 (0.58,7.2) Q4: 3.4 (0.97,11.7) p-trend = 0.04	MEHP Clinical loss OR (95% CI)	MEOHHP Clinical loss OR (95% CI)	DEHP Total loss Q2: 1.1 (0.61,2.0) Q3: 1.0 (0.56, 1.8) Q4: 1.6 (0.96, 2.7) p-trend = 0.06
Jukic et al. (2016); High; 221	6.6 (3.8–11)	30.3	DEHP T2: 0.32 (0.13, 0.77) * T3: 0.49 (0.19, 1.24)			
Yi et al. (2016); Low; 300	10.5 (mean)	6.2 (mean)		Continuous 1.78 (1.35,2.34) * Quartiles Q2: 3.3 Q3: 4.0 Q4: 7.2	Continuous 1.28 (0.84,1.95) Quartiles Q2: 1.6 Q3: 1.9 Q4: 4.6	
Toft et al. (2012); Low; 242	15.4 (7.5–27)	33.3	T2: 2.16 (0.63, 7.39) T3: 2.73 (0.78, 9.54) T1: 10.83 (1.16, 101) T3: 40.67 (4.48, 370)	T2: 0.17 (0.03, 0.95) * T3: 0.25 (0.05, 1.28)	T2: 0.55 (0.13, 2.29) T3: 0.55 (0.13, 2.35)	
Mu et al. (2015); Low; 304	20.6/4.8 (cases/non-cases) (1.5–103 cases)			Q2: 0.97 (0.47,1.98) Q3: 0.71 (0.34, 1.51) Q4: 1.30 (0.66, 2.55)		

* $p < 0.05$, results that support an association are shaded. Dark gray represents one or more of the following: $p < 0.05$, large effect size (e.g., OR $> = 1.5$), or exposure-response trend across categories of exposure. Studies are sorted by median exposure level. Q = quartile; T = tertile.

Table 11

Associations between DBP, DIBP, BBP, and DEP metabolite levels and spontaneous abortion

Reference; Study Confidence; N	Median exposure (IQR), (ng/mL)	Early loss OR (95% CI)	Clinical loss OR (95% CI)
DBP (measured by MBP)			
Messerlian et al. (2016); High; 256	12.6 (6.8–21)	Q2: 0.66 (0.22, 1.9) Q3: 0.78 (0.28, 2.2) Q4: 0.58 (0.19, 1.7)	"No notable association"
Yi et al. (2016); Low; 300	47 (mean)		1.00 (0.67, 1.50)
Jukic et al. (2016); High; 221	80 (49–127)	T2: 1.10 (0.47, 2.58) T3: 1.12 (0.46, 2.74)	Q2: 0.85 (0.44, 1.64) Q3: 0.82 (0.42, 1.59) Q4: 1.33 (0.72, 2.48)
Mu et al. (2015); Low; 304	86/44 (cases/noncases) (case IQR 10–240)		T2: 0.87 (0.21, 3.57) T3: 0.51 (0.12, 2.21)
Toft et al. (2012); Low; 242	174 (107–315)	T2: 1.25 (0.38, 4.09) T3: 1.64 (0.52, 5.20)	
DIBP (measured by MIBP)			
Jukic et al. (2016); High; 221	3.1 (1.9–5.0)	T2: 1.08 (0.47, 2.45) T3: 0.73 (0.32, 1.68)	
Messerlian et al. (2016); High; 256	7.2 (3.8–12)	Q2: 0.53 (0.17, 1.6) Q3: 1.2 (0.59, 2.5) Q4: 0.71 (0.26, 1.9)	"No notable association"
Yi et al. (2016); Low; 300	32 (mean)		0.85 (0.60, 1.19)
Mu et al. (2015); Low; 304	41/19 (cases/noncases) (case IQR 8.0–125)		Q2: 1.43 (0.71, 2.88) Q3: 1.38 (0.68, 2.79) Q4: 2.48 (1.28, 4.79)*
BBP (measured by MBzP)			
Messerlian et al. (2016); High; 256	3.2 (1.7–6.1)	Q2: 1.2 (0.46, 3.3) Q3: 1.2 (0.42, 3.2) Q4: 1.1 (0.39, 3.1)	"No notable association"
Toft et al. (2012); Low; 242	12 (7.3–23)	T2: 2.39 (0.70, 8.22) T3: 3.11 (0.87, 11.09)	T2: 1.08 (0.25, 4.66) T3: 0.96 (0.20, 4.59)
Jukic et al. (2016); High; 221	40 (25–69)	T2: 1.65 (0.69, 3.98) T3: 0.90 (0.36, 2.27)	
DEP (measured by MEP)			
Yi et al. (2016); Low; 300	12 (mean)		0.75 (0.57, 0.98)
Messerlian et al. (2016); High; 256	51 (21–133)	Q2: 0.96 (0.41, 2.30) Q3: 0.65 (0.23, 1.87) Q4: 0.86 (0.34, 2.16)	"No notable association"
Mu et al. (2015); Low; 304	57/16 (cases/noncases) (case IQR 3.8–322)		Q2: 1.04 (0.52, 2.08) Q3: 1.36 (0.70, 2.65) Q4: 1.88 (0.99, 3.58)
Jukic et al. (2016); High; 221	134 (73–296)	T2: 0.73 (0.31, 1.73) T3: 1.06 (0.48, 2.33)	NR
Toft et al. (2012); Low; 242	224 (113–419)	T2: 1.29 (0.43, 3.83) T3: 1.13 (0.36, 3.59)	T2: 2.19 (0.42, 11.45) T3: 4.63 (0.92, 23.26)

* $p < 0.05$, results that support an association are shaded. Dark gray represents one or more of the following: $p < 0.05$, large effect size (e.g., $OR > 1.5$), or exposure-response trend across categories of exposure. Light gray represents other supportive results. Studies within individual phthalates are sorted by median exposure level. Q = quartile; T = tertile.

Table 12

Epidemiology studies of preterm birth/gestational duration

Reference	Study description				Includes metabolites of:								Study evaluation				
	Population	Exposure	Outcome	DEH- P	DIN- P	DB- P	DIB- P	BB- P	DE- P	Exposur E	Outcome	Selection	Confounding	Analysis	Overall confidence		
Included Casas et al. (2016) <i>a</i>	Pregnancy cohort in Spain (N = 390)	Two urine samples at ~12 and 32 wk gestation	GD calculated from ultrasound and last menstrual period	✓	✓	✓	✓	✓	✓	G	G	G	G	G	High		
Ferguson et al. (2014a); Ferguson et al. (2014b); Chen et al. (2015);	Nested case-control within pregnancy cohort in U.S. (N= 130 cases preterm birth, 352 controls)	Up to four urine samples from menstrual period	GD calculated from ultrasound and last	✓	✓	✓	✓	✓	✓	G	A	A	A	G	High		
Gao et al. (2016) <i>a</i>	Pregnancy cohort in China (N= 3,103)	Single urine sample in 1 st trimester	GD calculated using last menstrual period or ultrasound	✓	✓	✓	✓	✓	✓	A/P	A	A	A	A	Medium		
Meeker et al. (2009)	Nested case-control in pregna cohort in Mexico (N = 30 cases and 30 controls)	Single urine sample in 3 rd trimester	GD calculated from last menstrual period (recall)	✓	✓	✓	✓	✓	✓	A/P	A	A	A	A	Medium		
Polanska et al. (2016)	Pregnancy cohort in Poland (N= 165)	Single urine sample in 3 rd trimester	GD estimated using last menstrual period or ultrasound	✓	✓	✓	✓	✓ ^b	✓	A/P	A	A	G	D	Medium		
Shoaff et al. (2016)	Pregnancy cohort in U.S. (N = 368)	Two urine samples at ~16 and 26 wk gestation	GD calculated from last menstrual period or other method	✓	✓	✓	✓	✓	✓	G	A	A	G	A	High		

Reference	Study description			Includes metabolites of:							Study evaluation					Overall confidence
	Population	Exposure	Outcome	DEH-P	DIN-P	DB-P	DIB-P	BB-P	DE-P	Exposur	Outcome	Selection	Confounding	Analysis		
Smarr et al. (2015) ^a	Preconception cohort in U.S. (n=233)	Single urine sample at baseline of trying to conceive	GD calculated from conception identified pro specively	✓	✓ ^b	✓	✓	✓	✓	D	G	A	A	A	Medium	
Watkins et al. (2016)	Pregnancy cohort in U.S. (N=68)	Two urine samples in 1 st trimester and at delivery	GD calculated using last menstrual period or ultrasound	✓	✓	✓	✓ ^c	✓	✓	G	A	A	A	D	Medium	
Total Studies per Phthalate																
				6	2	6	5	6	6	6						

Excluded studies: Exposure based on tissues other than urine: Brucker-Davis et al. (2010), Huang et al. (2009), Huang et al. (2014), Huang et al. (2003), Li et al. (2016). Exposure based on occupational exposure with unvalidated matrix: Burdorf et al. (2011). Study entry late in pregnancy (most or all enrollment in 3rd trimester): Adibi et al. (2009), Su et al. (2014), Suzuki et al. (2010), Weinberger et al. (2014), Whyatt et al. (2009), Wolff et al. (2008).

G = good; A = adequate; D = deficient; A/D = adequate for short chain phthalates, deficient for long chain phthalates; GD = gestational duration.

^a Additional analyses pertaining to preterm birth obtained from the study authors.

^b Considered *critically deficient* for this phthalate due to > 50% less than the LOD.

^c Results presented for summed DBP (MBP and MIBP), and included in analysis under DBP due to higher exposure levels.

Table 13

Evidence profile table for preterm birth and DEHP, DBP, and DEP.

Phthalate	Studies (design in parentheses)	Factors that increase confidence	Factors that decrease confidence	Summary of findings	Confidence judgment for outcome
DEHP	High confidence Casas et al. (2016) (C) Ferguson et al. (2014a) (CC) Shoaff et al. (2016) (C) Medium confidence Gao et al. (2016) (C) Meeker et al. (2009) (CC) Smarr et al., 2015 (C)	<ul style="list-style-type: none"> Minimal concerns for bias Exposure-response gradient within one study Biological plausibility 	<ul style="list-style-type: none"> Some unexplained inconsistency 	Increased odds of preterm birth with higher DEHP exposure in 4/6 studies (Casas et al., 2016; Ferguson et al., 2014b; Meeker et al., 2009; Smarr et al., 2015). Statistically significant exposure-response gradient observed in Ferguson et al. (2014b). Similar effect sizes among studies with elevated ORs. Biological plausibility - Ferguson et al. (2016) showed that oxidative stress may act as a mediator between phthalate exposure and preterm birth.	MODERATE
DBP	High confidence Casas et al. (2016) (C) Ferguson et al. (2014a) (CC) Shoaff et al. (2016) (C) Medium confidence Gao et al. (2016) (C) Meeker et al. (2009) (CC) Smarr et al., 2015 (C)	<ul style="list-style-type: none"> Minimal concerns for bias Large effect size in one study (OR = 4.5) Biological plausibility 	<ul style="list-style-type: none"> Some unexplained inconsistency 	Increased odds of preterm birth with higher DBP exposure in 3/6 studies (Casas et al., 2016; Ferguson et al., 2014b; Meeker et al., 2009), with statistical significant in Ferguson et al. (2014a, b) and Meeker et al. (2009). Biological plausibility - Ferguson et al. (2016) showed that oxidative stress may act as a mediator between phthalate exposure and preterm birth.	MODERATE
DEP	High confidence Casas et al. (2016) (C) Ferguson et al. (2014a) (CC) Shoaff et al. (2016) (C) Medium confidence Gao et al. (2016) (C) Meeker et al. (2009) (CC) Smarr et al., 2015 (C)	<ul style="list-style-type: none"> Minimal concerns for bias Lower study sensitivity (i.e., exposure levels) could account for remaining inconsistency Biological plausibility 		Increased odds of preterm birth with higher DEP exposure in 4/6 studies (Casas et al., 2016; Ferguson et al., 2014b; Meeker et al., 2009; Shoaff et al., 2016). Biological plausibility - Ferguson et al. (2016) showed that oxidative stress may act as a mediator between phthalate exposure and preterm birth.	MODERATE

C: cohort, CC: case-control, CS: cross-sectional, OR = odds ratio.