EPA Document #EPA-740-R-25-030
December 2025
Office of Chemical Safety and
Pollution Prevention

United States
Environmental Protection Agency

Meta-analysis and Benchmark Dose Modeling of Fetal Testicular Testosterone for Di(2-ethylhexyl) Phthalate (DEHP), Dibutyl Phthalate (DBP), Butyl Benzyl Phthalate (BBP), Diisobutyl Phthalate (DIBP), and Dicyclohexyl Phthalate (DCHP)

Technical Support Document for the Risk Evaluations

CASRNs: 117-81-7 (DEHP), 84-74-2 (DBP), 85-68-7 (BBP), 84-69-5 (DIBP), and 84-61-7 (DCHP)

December 2025

TABLE OF CONTENTS

LIST OF TABLES

LIST OF APPENDIX FIGURES

LIST OF APPENDIX TABLES

KEY ABBREVIATIONS AND ACRONYMS

AIC	Akaike information criterion
AGD	Anogenital distance
BBP	Butyl benzyl phthalate
BMD	Benchmark dose
BMDL	Benchmark dose (lower confidence limit)
BMR	Benchmark response
CASRN	Chemical abstracts service registry number
CRA	Cumulative risk assessment
DBP	Dibutyl phthalate
DCHP	Dicyclohexyl phthalate
DEHP	Di(2-ethylhexyl) phthalate
DIBP	Diisobutyl phthalate
DINP	Diisononyl phthalate
EPA	(U.S) Environmental Protection Agency (or "the Agency")
GD	Gestation day
MOA	Mode of action
NASEM	National Academies of Sciences, Engineering, and Medicine
NR	Nipple/areolae retention
OCSPP	Office of Chemical Safety and Pollution Prevention
OPPT	Office of Pollution Prevention and Toxics
RPF	Relative potency factor
SACC	Science Advisory Committee on Chemicals
SD	Sprague-Dawley (rat)
TSCA	Toxic Substances Control Act
UF	Uncertainty factor
U.S.	United States

ACKNOWLEDGEMENTS

Acknowledgements
The Assessment Team gratefully acknowledges the participation, input, and review comments from the U.S. Environmental Protection Agency (EPA or the Agency) Office of Pollution Prevention and Toxics (OPPT) and Office of Chemical Safety and Pollution Prevention (OCSPP) senior managers and science advisors, as well as intra-agency reviewers. The Agency is also grateful for assistance from EPA contractors SRC, Inc. (Contract No. 68HERH19D0022).

Special acknowledgement is given for the contributions of technical experts from EPA's Children's Health Protection Division, including Chris Brinkerhoff for providing review of this technical support document.

Docket
Supporting information can be found in the public dockets Docket IDs (EPA-HQ-OPPT-2018-0504, EPA-HQ-OPPT-2018-0434, EPA-HQ-OPPT-2018-0503, EPA-HQ-OPPT-2018-0433, and EPA-HQ-OPPT-2018-0501).

Disclaimer
Reference herein to any specific commercial products, process, or service by trade name, trademark, manufacturer, or otherwise does not constitute or imply its endorsement, recommendation, or favoring by the United States Government.

Author: Anthony Luz

Contributors: John Allran, Collin Beachum (Branch Supervisor), Brandall Ingle-Carlson, Ashley Peppriell, and Susanna Wegner

Technical Support: Mark Gibson, Hillary Hollinger, and S. Xiah Kragie

1 BACKGROUND

This technical support document (TSD) is for the risk evaluations for butyl benzyl phthalate (BBP) (U.S. EPA, 2025f), dibutyl phthalate (DBP) (U.S. EPA, 2025g), dicyclohexyl phthalate (DCHP) (U.S. EPA, 2025h), diethylhexyl phthalate (DEHP) (U.S. EPA, 2025i), diisobutyl phthalate (DIBP) (U.S. EPA, 2025j), as well as the *Technical Support Document for the Cumulative Risk Analysis of Di(2-ethylhexyl) Phthalate (DEHP), Dibutyl Phthalate (DBP), Butyl Benzyl Phthalate (BBP), Diisobutyl Phthalate (DIBP), Dicyclohexyl Phthalate (DCHP), and Diisononyl Phthalate (DINP) Under the Toxic Substances Control Act (TSCA)* (U.S. EPA, 2025m).

In 2017, the National Academies of Sciences, Engineering, and Medicine (NASEM) demonstrated the utility of a meta-analysis and meta-regression approach to combine fetal rat testicular testosterone data from multiple studies of similar design prior to conducting benchmark dose (BMD) modeling (NASEM, 2017). Meta-analysis is a statistical procedure that can be used to summarize outcomes from a number of studies and explore sources of heterogeneity in the data through use of random effects models. Therefore, meta-analysis can help overcome limitations associated with results from individual studies.

In the mode of action (MOA) for "phthalate syndrome," which has been described by EPA elsewhere (U.S. EPA, 2023), decreased fetal testicular testosterone is an early, upstream event in the MOA that precedes downstream apical outcomes such as male nipple retention, decreased anogenital distance, and male reproductive tract malformations (*e.g.,* hypospadias, cryptorchidism). Decreased fetal testicular testosterone should occur at doses that are lower than or equal to doses that cause downstream apical outcomes associated with a disruption of androgen action. Therefore, consistent with the best available science, EPA conducted an updated meta-analysis and BMD modeling analysis of decreased fetal rat testicular testosterone using similar methods as employed by NASEM (2017) and incorporating more recent studies. The purpose of this updated meta-analysis and BMD modeling analysis is to provide the most up-to-date dose-response information in support of the individual phthalate risk evaluations as well as the cumulative risk assessment of phthalates. The remainder of this TSD is organized as follows:

- Section 2 provides an overview of the methods employed by EPA for the updated meta-analysis and BMD modeling analysis of fetal rat testicular testosterone. A description of differences between the NASEM (2017) analysis and EPA's updated analysis is also provided.
- Section 3 summarizes the results of EPA's replicate analysis of NASEM's meta-analysis and BMD modeling analysis of DIBP.
- Section 5 summarizes EPA's updated meta-analysis and BMD modeling results of fetal rat testicular testosterone for DBP (Section 5.1), DEHP (Section 5.2), DIBP (Section 5.3), BBP (Section 5.4), and DCHP (Section 5.5).
- Section 6 compares BMD modeling results obtained by EPA as part of the updated analysis and results from NASEM (2017).
- Section 7 describes EPA's conclusions.

2 METHODS

In 2017, NASEM demonstrated the utility of meta-analysis and meta-regression to summarize several outcomes from experimental animal studies (NASEM, 2017). The 2017 NASEM analysis included reduced fetal testicular testosterone, reduced male anogenital distance (AGD), and increased incidence of hypospadias in rodents following oral exposure to DEHP, DBP, BBP, DIBP, and DINP. DCHP was not included as part of the NASEM analysis. Boxes 3-3 and 3-4 in (NASEM, 2017) provide detailed descriptions of the meta-analysis approach employed by NASEM. Briefly, NASEM conducted meta-analyses using the Metafor (Version 2.0.0) meta-analysis package for R, which employs a standard random effects model using the Restricted Maximum Likelihood Estimate. The meta-analyses conducted by NASEM focused on the dose-response relationship and employed three models, linear, log-linear, and linear-quadratic models. The linear meta-regressions with dose in original and log-transformed units were used to assess the presence or absence of a gradient. For the linear and linear-quadratic models, BMD values were estimated based on benchmark response (BMR) levels of 5 and 40 percent. NASEM did not provide explicit justification for selection of a BMR of 5 percent. However, justification for the BMR of 5 percent can be found elsewhere (U.S. EPA, 2012; Allen et al., 1994a, b; Faustman et al., 1994).

As discussed in EPA's *Benchmark Dose Technical Guidance* (U.S. EPA, 2012), a BMR of 5 percent is supported for BMD modeling of most endpoints in developmental and reproductive studies. Comparative analyses of a large database of developmental toxicity studies demonstrated that developmental NOAELs are approximately equal to the $BMDL_5$ (Allen et al., 1994a, b; Faustman et al., 1994). NASEM (2017) also modeled a BMR of 40 percent using the following justification: "previous studies have shown that reproductive-tract malformations were seen in male rats when fetal testosterone production was reduced by about 40% (Gray et al., 2016; Howdeshell et al., 2015)." The R code used by NASEM to conduct all meta-analyses is publicly available (https://github.com/wachiuphd/NASEM-2017-Endocrine-Low-Dose; accessed December 16, 2025).

As part of its updated analysis, EPA used a similar meta-analysis and BMD modeling approach as employed by NASEM (2017), but with several notable differences. First, EPA used the most recent version of the R Metafor package (Version 4.6.0) available at the time of the updated analysis, while NASEM used Metafor Version 2.0.0. However, EPA also conducted the updated analysis with Metafor Version 2.0.0 so that results from the two different versions of Metafor could be compared. Similar to the NASEM approach, EPA's updated meta-analysis focused on the dose-response relationship and employed the linear and log-linear models for trend analysis and the linear and linear-quadratic models for BMD analysis. Another notable difference between the NASEM analysis and EPA's updated analysis is that EPA evaluated BMRs of 5, 10, and 40 percent, while NASEM evaluated BMRs of 5 and 40 percent. EPA added evaluation of a BMR of 10 percent because BMD modeling of fetal testosterone conducted by NASEM (2017) indicated that BMD_5 estimates are more than three-fold below the lowest dose with empirical testosterone data for several of the phthalates (*e.g.*, DIBP). As discussed in EPA's *Benchmark Dose Technical Guidance* (U.S. EPA, 2012) "For some datasets the observations may correspond to response levels far in excess of a selected BMR and extrapolation sufficiently below the observable range may be too uncertain to reliably estimate BMDs/BMDLs for the selected BMR." Therefore, EPA modeled a BMR of 10 percent because datasets for some of the phthalates may not include sufficiently low doses to support modeling of a 5 percent response level. For the linear and linear-quadratic models, BMD values were estimated based on BMR levels of 5, 10, and 40 percent. The linear meta-regressions with dose in original and log-transformed units were used to assess the presence or absence of a gradient. BMD models were examined for a visual fit to the data, and the best-fit model was determined based on the lowest Akaike information criterion (AIC).

One additional difference between the NASEM (2017) analysis and EPA's updated analysis is that NASEM included an analysis in which rat data were subjected to a subgroup analysis by strain because of potential differential sensitivity across strains. NASEM conducted this subgroup analysis only for DEHP. EPA did not include a subgroup analysis as part of its updated meta-analysis and BMD modeling analysis because (1) the number of new studies identified by EPA evaluating fetal testicular testosterone is small; (2) none of the new studies provide obviously different results from the studies analyzed by NASEM; and (3) only studies of Sprague-Dawley rats are available for DIBP, BBP, and DCHP. Further, NASEM only identified slight differences in strain sensitivity for effects on fetal testicular testosterone for DEHP (with Sprague-Dawley rats being slightly more sensitive than Wistar); however, the apparent difference in sensitivity appears to be due to model choice—instead of a true difference in strain sensitivity. For example, the linear model provided the best fit (based on lowest AIC) for Wistar rats, while the Linear-Quadratic Model provided the best fit for Sprague-Dawley and the analysis of all strains combined.

As part of the updated meta-analysis, EPA utilized all of the same fetal rat testicular testosterone data included in the original NASEM (2017) analysis, as well as new fetal rat testosterone data identified through the 2019 TSCA literature searches for DBP, DEHP, DIBP, BBP, and DCHP. EPA also considered new literature identified outside of the 2019 TSCA literature searches that was identified through the literature searches conducted in 2022 in support of EPA's *Draft Proposed Approach for Cumulative Risk Assessment of High-Priority Phthalates and a Manufacturer-Requested Phthalate under the Toxic Substances Control Act* (U.S. EPA, 2023).

Consistent with the meta-analysis and BMD modeling approach employed by NASEM (2017), new fetal rat testicular testosterone data were only included in the updated meta-analysis if the following criteria were met:

- Study conducted with pregnant rats (all strains considered relevant, including Sprague-Dawley, Wistar, Long Evans, F344, etc.). For the updated analysis, studies of mice were excluded because rats are considered the more sensitive species.
- Study exposed rats via the oral route.
- Study measured fetal testis testosterone content or *ex vivo* fetal testicular testosterone production. Studies measuring only serum or plasma testosterone were excluded. Studies measuring testosterone at non-fetal lifestages were excluded. Studies measuring testosterone production following stimulation with luteinizing hormone were excluded.
- Studies measuring testosterone levels within fetal life but outside of the male programming window (defined by NASEM as gestational days (GD) 16–18) were included because fetal Leydig cell testosterone production sensitivity to phthalate exposure encompasses the entirety of fetal life when the testis is producing testosterone.
- Study fully reported data (*i.e.*, mean, standard deviation or standard error, and sample size) to support extraction and inclusion in meta-analysis. Note: when new fetal testicular testosterone data were presented graphically only, and not in a tabular form, EPA did not extract the data and did not include the data in its updated analysis.

As will described further in Section 5, EPA identified new fetal testicular testosterone data for DEHP, DBP, DIBP, BBP, and DCHP to support the updated meta-analysis. All studies included in the updated meta-analysis and BMD modeling analysis of fetal testicular testosterone were evaluated for study quality as described in the systematic review protocols for DCHP (U.S. EPA, 2025l).

3 OVERVIEW OF SACC RECOMMENDATIONS

This technical support document was released in draft for public comment and was peer-reviewed by the Science Advisory Committee on Chemicals (SACC) during the August 4 to 8, 2025 peer-review meeting (U.S. EPA, 2025k). SACC provided EPA with several recommendations, including 1) to explore additional tools and methods for BMD modeling and meta-analysis to address several limitations and uncertainties associated with the use of Metafor and 2) to consider additional analyses to determine if phthalate dose-response curves are parallel. These recommendations are discussed further below, along with a brief description of how they were addressed by EPA. Readers are directed to EPA's response to public comments summary document and EPA's response to the 2025 phthalates SACC meeting report for further details.

1. SACC noted that Metafor includes two models, including linear and linear-quadratic models, which might not have the ability to fit sigmoidal shape testosterone dose-response curves. SACC recommended EPA consider using EPA's current BMD software (BMDS), which contains a wider suite of models (*i.e.*, Exponential, Hill, Polynomial, Power, Linear models), to address this uncertainty. In response, EPA conducted additional BMD modeling of individual fetal testicular testosterone datasets for DBP (index chemical), DIBP, BBP, and DCHP. Results from this additional BMD modeling is discussed in the individual non-cancer human health hazard assessments for DBP (U.S. EPA, 2025b), DIBP (U.S. EPA, 2025d), BBP (U.S. EPA, 2025a), and DCHP (U.S. EPA, 2025c), as well as the Technical Support Document for the Cumulative Risk Analysis of DEHP, DBP, BBP, DIBP, DCHP, and DINP Under TSCA (U.S. EPA, 2025m). As discussed further in these documents, BMD modeling of individual fetal testicular testosterone datasets using EPA's BMD software provided similar results to Metafor. This indicates that models in Metafor provide reasonable BMD/BMDL estimates, and support EPA's use of Metafor for meta-analysis and BMD modeling.
2. SACC noted that Metafor Version 4.6.0 did not allow BMD_5 values or relative potency factors (RPFs) to be estimated for BBP or DIBP, while the older version of Metafor (Version 2.0.0) allowed for BMD_5 estimates for all phthalates included in the cumulative assessment. SACC recommended EPA consider use of older Metafor Version 2.0.0 results to calculate RPFs, since this version of Metafor allowed BMD_5 estimates to be derived for all phthalates included in the cumulative assessment. As discussed in the Technical Support Document for the Cumulative Risk Analysis of DEHP, DBP, BBP, DIBP, DCHP, and DINP Under TSCA (U.S. EPA, 2025m), EPA calculated candidate RPFs using BMD_5 estimates from Metafor Version 2.0.0 and compared these RPFs to the selected RPFs based on BMD_{40} estimates from Metafor Version 4.6.0. Overall, RPFs calculated at both response levels using different Versions of Metafor are similar. For example, the selected RPF for DEHP is 0.84 (Metafor Version 4.6.0) compared to an RPF of 0.88 (Version 2.0.0) (4.8% difference); the selected RPF for DIBP is 0.53 (Version 4.6.0) compared to an RPF of 0.42 (Version 2.0.0) (21% difference); the selected RPF for BBP is 0.52 (Version 4.6.0) compared to an RPF of 0.48 (Version 2.0.0) (7.7% difference); the selected RPF for DCHP is 1.66 (Version 4.6.0) compared to and RPF of 1.83 (Version 2.0.0) (10% difference); and the selected RPF for DINP is 0.21 (Version 4.6.0) compared to an RPF of 0.19 (Version 2.0.0) (9.5% difference). The fact the selected RPFs based on BMD_{40} estimates calculated using Metafor Version 4.6.0 are similar to RPFs based on BMD_5 estimates calculated using Metafor Version 2.0.0 further increases EPA's confidence in the selected RPFs calculated using Metafor Version 4.6.0.
3. SACC recommended that EPA address one public comment that indicates that Bayesian Hierarchical Modeling represents the state of the science for deriving BMD estimates and RPFs for phthalates. As discussed further in the Technical Support Document for the Cumulative Risk

Analysis of DEHP, DBP, BBP, DIBP, DCHP, and DINP Under TSCA (U.S. EPA, 2025m), EPA considered the suggested Bayesian Hierarchical BMD modeling approach. EPA recognizes that although the Bayesian Hierarchical Modeling approach may represent an alternative method to estimate BMD values and RPFs, the new method is not yet available as open-source software and was not reasonably available to EPA. Importantly, EPA considers its current analysis using Metafor to be scientifically valid and appropriate for deriving BMD estimates and RPFs. This is because candidate RPFs estimated using Metafor did not vary significantly across response levels providing evidence of parallel dose-response curves. Further, BMD/BMDL estimates derived using the Metafor approach and EPA's BMD software provided similar results.

4. Although SACC recognized that parallel dose-response curves are not required for application of the RPF approach, SACC stated that demonstration of parallel curves might increase confidence in EPA's cumulative risk assessment approach and recommended EPA attempt to address this uncertainty. As discussed further in the Technical Support Document for the Cumulative Risk Analysis of DEHP, DBP, BBP, DIBP, DCHP, and DINP Under TSCA (U.S. EPA, 2025m), candidate RPF values did not vary significantly across the 5, 10, or 40 percent response levels for DEHP, DBP, DINP, or DCHP, or across the 10 and 40 percent response levels for DIBP, indicating parallel dose-response curves for these phthalates.

Overall, EPA's fundamental approach to dose-response assessment of fetal testicular testosterone data (*i.e.*, meta-analysis and BMD analysis using Metafor) did not change after taking into consideration SACC recommendations and public comments.

4 REPLICATION OF NASEM META-ANALYSIS AND BENCHMARK DOSE MODELING APPROACH

As a proof of principle and to demonstrate replicability of NASEM's meta-analysis and BMD modeling approach, EPA first used publicly available R-code provided by NASEM to attempt to replicate results from the 2017 NASEM meta-analysis and BMD modeling analysis of fetal testicular testosterone in rats for DIBP. The analysis by NASEM (2017) included *ex vivo* fetal testicular testosterone production data from two rat studies of DIBP (Hannas et al., 2011; Howdeshell et al., 2008). EPA used the same *ex vivo* fetal testicular testosterone production data from these two studies as part of its replicate analysis.

Initially, EPA was unable to replicate the meta-analysis and BMD modeling results reported by NASEM (2017) for DIBP, with results varying significantly between the NASEM and EPA's analysis (Table 4-1 and Table 4-2). The Agency determined the discrepancies between the results obtained by NASEM (2017) and its replicate analysis were due to updates in the Metafor package in R. In 2017, the NASEM analysis relied on Metafor Version 2.0.0. EPA was able to replicate the NASEM (2017) results for DIBP exactly using Metafor Version 2.0.0 (Table 4-1 and Table 4-2). However, use of Metafor version 4.6.0 resulted in different meta-analysis and BMD modeling results for DIBP (Table 4-1 and Table 4-2). EPA was unable to determine the precise reasons for the deviations in the results using Metafor Versions 2.0.0 and 4.6.0. The primary functions from Metafor used in the meta-analysis repeatedly are rma() and forest(), which have many updates in each version of Metafor. The complete Metafor package changelog is available at https://wviechtb.github.io/metafor/news/index.html (accessed December 16, 2025).

Table 4-1 and Table 4-2 provide a comparison of overall meta-analysis results and BMD modeling results, respectively, obtained by NASEM (2017) and by EPA using Metafor Versions 2.0.0 and 4.6.0. Additional meta-analysis results (*i.e.*, forest plots) and BMD model fit curves obtained by EPA using Metafor Versions 2.0.0 and 4.6.0 are provided in Appendix A.1. As can be seen from Table 4-2, for NASEM (2017) and EPA's analysis using Metafor Version 2.0.0, there was a statistically significant overall effect and linear trends in $\log_{10}$(dose) and dose and both analyses support BMD_5 and BMD_{40} values of 27 mg/kg-day (95% confidence interval [95% CI]: 23, 34) and 271 mg/kg-day (95% CI: 225, 342), respectively, based on the best fit linear model (based on lower AIC than the linear quadratic model). EPA's analysis using Metafor Version 4.6.0 provided nearly identical results as Metafor Version 2.0.0 for the linear model (Table 4-2). However, using Metafor Version 4.6.0 the linear-quadratic model provided the best fit (based on lowest AIC) and supports a BMD_{40} of 263 mg/kg-day. A BMD_5 could not be derived using Metafor Version 4.6.0 for the linear-quadratic model.

Overall, EPA selected BMD modeling results obtained using Metafor Version 4.6.0 for use in the single phthalate risk evaluations and phthalate cumulative risk assessment because these results were obtained using the most up-to-date version of the Metafor package available at the time of the updated meta-analysis and BMD modeling analysis. However, EPA conducted all subsequent meta-analyses and BMD modeling analyses reported in Section 5 using both versions of Metafor (version 2.0.0 and version 4.6.0) so that results could be compared.

Table 4-1. Replication of NASEM (2017) Results: Comparison of Overall Meta-Analyses of Rat Studies of DIBP and Fetal Testicular Testosterone Using Metafor Version 2.0.0 and Version 4.6.0

Analysis	Estimate	Beta	CI, Lower Bound	CI, Upper Bound	P value	Tau	I^2	P value for Heterogeneity	AIC
NASEM (2017) analysis using Metafor Version 2.0.0 (from Table C6-11 in NASEM (2017))									
Overall	intercept	−82.31	−135.11	−29.52	0.002	71.76	96.96	0.000[a]	87.28
Trend in log10(dose)	log10(dose)	−169.23	−234.13	−104.33	0.000[a]	28.14	77.83	0.001	78.52
Linear in dose100	dose100	−18.84	−22.73	−14.94	0.000[a]	18.64	78.78	0.001	75.51*
Linear Quadratic in dose100	dose100	−11.61	−22.13	−1.08	0.031	12.22	57.12	0.02	77.04
Linear Quadratic in dose100	I(dose100^2)	−1.00	−2.42	0.42	0.169				
EPA analysis using Metafor Version 2.0.0									
Overall	intercept	−82.31	−135.11	−29.52	0.002	71.76	96.96	0.000[a]	87.28
Trend in log10(dose)	log10(dose)	−169.23	−234.13	−104.33	0.000[a]	28.14	77.83	0.001	78.52
Linear in dose100	dose100	−18.84	−22.73	−14.94	0.000[a]	18.64	78.78	0.001	75.51*
Linear Quadratic in dose100	dose100	−11.61	−22.13	−1.08	0.031	12.22	57.12	0.02	77.04
Linear Quadratic in dose100	I(dose100^2)	−1.00	−2.42	0.42	0.169				
EPA analysis using Metafor Version 4.6.0									
Overall	intercept	−82.31	−135.11	−29.52	0.00[a]	71.76	96.96	0.000[a]	87.28
Trend in log10(dose)	log10(dose)	−169.3	−234.13	−104.33	0.00[a]	28.14	77.83	0.001	78.52
Linear in dose100	dose100	−18.64	−27.52	−9.76	0.00[a]	65.25	97.85	0.00[a]	81.28
Linear Quadratic in dose100	dose100	−19.78	−50.04	10.48	0.20	54.97	96.42	0.00[a]	80.73*
Linear Quadratic in dose100	I(dose100^2)	0.14	−3.72	4.00	0.94				

* Indicates model with lowest Akaike information criterion (AIC).

Abbreviations: CI = confidence interval; I^2 = describes the percentage of the variability in effect estimates that is due to heterogeneity rather than sampling error; Tau = estimated standard deviation of the true underlying effect sizes across studies in the random-effects model meta-analysis

[a] p-value too small to calculate and rounded to zero.

Table 4-2. Replication of NASEM (2017) Results: Comparison of Benchmark Dose Estimates for Decreased Fetal Testicular Testosterone in Rats Following Gestational Exposure to DIBP using Metafor Version 2.0.0 and Version 4.6.0

Analysis	BMR	BMD (mg/kg-day)	CI, Lower Bound (mg/kg-day)	CI, Upper Bound (mg/kg-day)	AIC
NASEM (2017) analysis using Metafor Version 2.0.0 (from Tables C6-11 and C6-12 in NASEM (2017)) [a]					
Linear in dose100*	5%	27	23	34	75.51*
Linear in dose100*	40%	271	225	342	
Linear Quadratic in dose100	5%	43	23	127	77.04
Linear Quadratic in dose100	40%	341	239	453	
EPA analysis using Metafor Version 2.0.0[b]					
Linear in dose100*	5%	27	23	34	75.51*
Linear in dose100*	40%	271	225	342	
Linear Quadratic in dose100	5%	43	23	127	77.04
Linear Quadratic in dose100	40%	341	239	453	
EPA analysis using Metafor Version 4.6.0					
Linear in dose100	5%	28	19	53	81.28
Linear in dose100	40%	274	186	523	
Linear Quadratic in dose100*	5%	NA	NA	343	80.73*
Linear Quadratic in dose100*	40%	263	NA	585	

* Indicates model with lowest Akaike information criterion (AIC).

Abbreviations: BMD = benchmark dose; BMR = benchmark response; CI = confidence interval

[a] EPA noted an apparent discrepancy in the NASEM (2017) report. In Table 3-26, NASEM notes that no BMD/BMDL estimates could be generated at the 5% response level for DIBP because "the 5% change was well below the range of the data, but it will be 10 times lower because a linear model was used." However, in Table C6-12 of the NASEM report, BMD/BMDL estimates at the 5% response level are provided for DIBP for the best-fit linear model. In EPA's replicate analysis, identical BMD/BMDL estimates for the 5% response level were obtained. Therefore, BMD/BMDL estimates at the 5% response level for DIBP are reported in this table.

5 META-ANALYSIS AND BMD MODELING OF FETAL TESTICULAR TESTOSTERONE

5.1 Dibutyl Phthalate (DBP)[1]

EPA identified 29 studies of DBP evaluating testosterone (Table_Apx B-1). Of these studies, 8 met the criteria outlined in Section 2 for inclusion in the meta-analysis (Table 5-1). Seven of the eight studies evaluating fetal rat testicular testosterone content and/or *ex vivo* testosterone production were included in the 2017 NASEM meta-analysis. EPA identified new fetal rat testicular testosterone data from one study (Gray et al., 2021), which was included as part of the updated meta-analysis and BMD modeling analysis for DBP. Table 5-1 provides an overview of the eight studies included in the updated meta-analysis.

Twenty-one studies did not meet the inclusion criteria outlined in Section 2 and were excluded from EPA's updated meta-analysis for various reasons, as outlined in Table_Apx B-1. Of the 21 excluded studies, 5 were excluded from the original meta-analysis conducted by NASEM in 2017 due to data reporting issues (*e.g.*, N reported as range, not exact value or variance type (SEM, SD) not reported) (Li et al., 2015; van den Driesche et al., 2012; Clewell et al., 2009; Mahood et al., 2007; Lehmann et al., 2004). EPA excluded another six studies due to similar data reporting issues (*e.g.*, N reported as range, not exact value and/or data reported graphically only) (Spade et al., 2018; MacLeod et al., 2010; Drake et al., 2009; Howdeshell et al., 2007; Wilson et al., 2004; Mylchreest et al., 2002). Five studies were excluded because they evaluated serum (not testicular) testosterone during a postnatal (not fetal) lifestage (Ahmad et al., 2014; Giribabu et al., 2014; Kim et al., 2010; Scarano et al., 2010; Xiao-Feng et al., 2009). The last five studies were excluded because they evaluated testosterone in a species other than the rat (*i.e.*, mouse, rabbit, or monkey) (Li et al., 2023; Moody et al., 2013; McKinnell et al., 2009; Gaido et al., 2007; Higuchi et al., 2003).

For the eight included studies, EPA conducted the updated meta-analysis using random effects models, as implemented in the R Metafor package. Metafor versions 2.0.0 and 4.6.0 were used so that results could be compared. Additionally, the updated analysis included a sensitivity analysis to determine if the meta-analysis was sensitive to leaving out results from individual studies.

Table 5-1. Summary of Studies Included in EPA's Meta-analysis and BMD Modeling Analysis for DBP

Reference (TSCA Study Quality Rating)	Included in NASEM Meta-analysis and BMD Modeling Analysis?	Brief Study Description	Measured Outcome
(Martino-Andrade et al., 2008) (Medium)	Yes	Pregnant Wistar rats (7–8 dams/group) gavaged with 0, 100, 500 mg/kg-day DBP on GD 13–21	Fetal testis testosterone content on GD 21

[1] In addition to the meta-analysis, EPA also conducted additional BMD modeling of all individual studies of DBP in Table 5-1 reporting reduced fetal testicular testosterone using all standard continuous models in EPA's BMD software (BMDS Online Version 25.1). These BMD model results are reported in EPA's *Non-Cancer Human Health Hazard Assessment for Dibutyl Phthalate (DBP)* (U.S. EPA, 2025b).

Reference (TSCA Study Quality Rating)	Included in NASEM Meta-analysis and BMD Modeling Analysis?	Brief Study Description	Measured Outcome
(Furr et al., 2014) (High)	Yes	Pregnant SD rats (2–3 dams/group) gavaged with 0, 33, 50, 100, 300 mg/kg-day DBP on GD 14–18 (Block 18)	*Ex vivo* fetal testicular testosterone production (3-hour incubation) on GD 18
	Yes	Pregnant SD rats (3–4 dams/group) gavaged with 0, 1, 10, 100 mg/kg-day DBP on GD 14–18 (Block 22)	*Ex vivo* fetal testicular testosterone production (3-hour incubation) on GD 18
	Yes	Pregnant SD rats (3–4 dams/group) gavaged with 0, 1, 10, 100 mg/kg-day DBP on GD 14–18 (Block 26)	*Ex vivo* fetal testicular testosterone production (3-hour incubation) on GD 18
(Howdeshell et al., 2008) (High)	Yes	Pregnant SD rats (3–4 dams/group) gavaged with 0, 33, 50, 100, 300, 600 mg/kg-day DBP on GD 8–18	*Ex vivo* fetal testicular testosterone production (2-hour incubation) on GD 18
(Kuhl et al., 2007) (Low)	Yes	Pregnant SD rats (3–4 dams/group) gavaged with 0, 100, 500 mg/kg-day DBP on GD 18	Fetal testis testosterone content on GD 19
(Struve et al., 2009) (Medium)	Yes	Pregnant SD rats (7–9 dams/group) gavaged with 0, 112, 581 mg/kg-day DBP on GD 12–19	Fetal testis testosterone content on GD 19 (4-hour post-exposure)
		Pregnant SD rats (7–9 dams/group) gavaged with 0, 112, 581 mg/kg-day DBP on GD 12–19	Fetal testis testosterone content on GD 20 (24-hour post-exposure)
(Johnson et al., 2011) (Medium)	Yes	Pregnant SD rats (5–6 dams/group) gavaged with 0, 100 mg/kg-day DBP on GD 12–20	Fetal testis testosterone content on GD 20
		Pregnant SD rats (5–6 dams/group) gavaged with 0, 500 mg/kg-day DBP on GD 12–20	Fetal testis testosterone content on GD 20
(Johnson et al., 2007) (Medium)	Yes	Pregnant SD rats (5 dams/group) gavaged with 0, 1, 10, 100 mg/kg-day DBP on GD 19	Fetal testis testosterone content on GD 19
(Gray et al., 2021) (High)	No (new study)	Pregnant SD rats (3–4 dams/group) gavaged with 0, 300, 600, 900 mg/kg-day DBP on GD 14–18 (Block 70)	*Ex vivo* fetal testicular testosterone production (3-hour incubation) on GD 18

Reference (TSCA Study Quality Rating)	Included in NASEM Meta-analysis and BMD Modeling Analysis?	Brief Study Description	Measured Outcome
	No (new study)	Pregnant SD rats (3–4 dams/group) gavaged with 0, 300, 600, 900 mg/kg-day DBP on GD 14–18 (Block 71)	*Ex vivo* fetal testicular testosterone production (3-hour incubation) on GD 18

Overall meta-analyses and sensitivity analyses results obtained using Metafor Versions 2.0.0 and 4.6.0 are shown in Table 5-2 and Table 5-3, respectively. A comparison of BMD estimates obtained by NASEM (2017) and as part of EPA's updated analysis are shown in Table 5-4. Additional meta-analysis results (*i.e.*, forest plots) and BMD model fit curves are shown in Appendix A.2. For meta-analyses conducted using both versions of Metafor, there was a statistically significant overall effect and linear trends in $\log_{10}$(dose) and dose, with an overall effect that is large in magnitude (>50% change). For both meta-analyses, there was substantial, statistically significant heterogeneity in all cases ($I^2 > 80\%$ for Metafor v.2.0.0; $I^2 > 88\%$ for Metafor v.4.6.0). The statistical significance of these effects was robust to leaving out individual studies for analyses conducted with both versions of Metafor. Although there was substantial heterogeneity, standard deviation of the random effect (tau) was less than the estimated size of the effect at higher doses. Therefore, the heterogeneity does not alter the conclusion that gestational exposure to DBP reduces fetal testicular testosterone in the rat.

For meta-analyses conducted using both versions of Metafor, the linear-quadratic model provided the best fit (*i.e.*, had lower AIC than the linear model) (Table 5-4). BMD estimates from the linear-quadratic model were 15 mg/kg-day (95% CI: 11, 21) for a 5 percent change (BMR = 5%), 30 mg/kg-day (95% CI: 23, 43) for a 10 percent change (BMR = 10%), and 154 mg/kg-day (95% CI: 119, 211) for a 40 percent change (BMR = 40%) when Metafor Version 2.0.0 was used for the updated analysis including the new study by Gray et al. (2021). Similarly, BMD estimates from the linear-quadratic model were 14 mg/kg-day (95% CI: 9, 27) for a 5 percent change (BMR = 5%), 29 mg/kg-day (95% CI: 20, 54) for a 10 percent change (BMR = 10%), and 149 mg/kg-day (95% CI:101, 247) for a 40 percent change (BMR = 40%) when Metafor Version 4.6.0 was used to model all of the studies including the new data.

Notably, Metafor versions 2.0.0 and 4.6.0 provided similar BMD_5 (15 vs. 14 mg/kg-day), BMD_{10} (30 vs. 29 mg/kg-day), and BMD_{40} (154 vs. 149 mg/kg-day) estimates for the best fitting, linear-quadratic model (Table 5-4) for the updated analysis including the new study by Gray et al. (2021), and these results are similar to those obtained in the 2017 NASEM meta-analysis (*i.e.*, BMD_5 and BMD_{40} estimates of 12 and 125 mg/kg-day, respectively, based on the best fitting linear quadratic model). At the evaluated BMRs of 5 and 40 percent, inclusion of the new data results in slightly higher BMD_5 and BMD_{40} estimates with similar 95 percent confidence intervals compared to results obtained in the 2017 NASEM analysis.

Table 5-2. Updated Overall Meta-analyses and Sensitivity Analyses of Rat Studies of DBP and Fetal Testosterone (Metafor Version 2.0.0)

Analysis	Estimate	Beta	CI, Lower Bound	CI, Upper Bound	P value	Tau	I^2	P Value for Heterogeneity	AIC
Primary analysis									
Overall	intercept	−71.85	−95.76	−47.95	3.82E−09	67.01	95.60	2.74E-152	383.39
Trend in log10(dose)	log10(dose)	−62.44	−81.70	−43.19	2.08E−10	41.61	88.70	4.43E−50	349.26
Linear in dose100	dose100	−25.02	−28.72	−21.32	3.76E−40	32.26	83.67	2.85E−39	344.58
Linear Quadratic in dose100	dose100	−35.58	−46.64	−24.52	2.84E−10	30.36	80.93	7.99E−22	334.19*
Linear Quadratic in dose100	I(dose100^2)	1.61	0.02	3.19	4.73E−02	30.36	80.93	7.99E−22	334.19
Sensitivity analysis									
Overall minus Furr et al. (2014)	intercept	−88.38	−117.31	−59.45	2.14E−09	67.21	93.19	2.16E−55	270.22
Overall minus Johnson et al. (2007)	Intercept	−76.78	−102.25	−51.31	3.47E−09	68.66	96.10	3.84E−153	350.04
Overall minus Howdeshell et al. (2008)	intercept	−78.30	−105.70	−50.91	2.11E−08	70.83	95.72	3.63E−139	329.10
Overall minus Johnson et al. (2011)	intercept	−69.59	−93.70	−45.48	1.53E−08	65.39	95.51	3.39E−148	359.45
Overall minus Kuhl et al. (2007)	intercept	−72.06	−97.37	−46.75	2.39E−08	68.92	95.94	3.87E−152	362.13
Overall minus Martino-Andrade et al. (2008)	intercept	−72.43	−97.80	−47.06	2.19E−08	69.11	95.94	1.74E−152	362.26
Overall minus Struve et al. (2009)	intercept	−63.19	−86.77	−39.61	1.50E−07	62.87	95.50	2.53E−148	329.62
Overall minus Gray et al. (2021)	intercept	−56.97	−80.64	−33.31	2.37E−06	59.25	94.78	3.05E−115	311.44

* Indicates model with lowest Akaike information criterion (AIC).
Abbreviations: CI = confidence interval; I^2 = describes the percentage of the variability in effect estimates that is due to heterogeneity rather than sampling error; Tau = estimated standard deviation of the true underlying effect sizes across studies in the random-effects model meta-analysis

Table 5-3. Updated Overall Meta-analyses and Sensitivity Analyses of Rat Studies of DBP and Fetal Testosterone (Metafor Version 4.6.0)

Analysis	Estimate	Beta	CI, Lower Bound	CI, Upper Bound	P value	Tau	I^2	P Value for Heterogeneity	AIC
Primary analysis									
Overall	intercept	−71.85	−95.76	−47.95	3.82E−09	67.01	95.60	2.74E−152	383.39
Trend in log10(dose)	log10(dose)	−62.44	−81.70	−43.19	2.08E−10	41.61	88.70	4.43E−50	349.26
Linear in dose100	dose100	−25.69	−31.55	−19.83	8.64E−18	57.78	94.26	3.38E−119	354.71
Linear Quadratic in dose100	dose100	−36.78	−54.53	−19.03	4.89E−05	54.79	93.26	1.72E−117	343.82*
Linear Quadratic in dose100	I(dose100^2)	1.70	−0.86	4.26	1.94E−01	54.79	93.26	1.72E−117	343.82
Sensitivity analysis									
Overall minus Furr et al. (2014)	intercept	−88.38	−117.31	−59.45	2.14E−09	67.21	93.19	2.16E−55	270.22
Overall minus Johnson et al. (2007)	intercept	−76.78	−102.25	−51.31	3.47E−09	68.66	96.10	3.84E−153	350.04
Overall minus Howdeshell et al. (2008)	intercept	−78.30	−105.70	−50.91	2.11E−08	70.83	95.72	3.63E−139	329.10
Overall minus Johnson et al. (2011)	intercept	−69.59	−93.70	−45.48	1.53E−08	65.39	95.51	3.39E−148	359.45
Overall minus Kuhl et al. (2007)	intercept	−72.06	−97.37	−46.75	2.39E−08	68.92	95.94	3.87E−152	362.13
Overall minus Martino-Andrade et al. (2008)	intercept	−72.43	−97.80	−47.06	2.19E−08	69.11	95.94	1.74E−152	362.26
Overall minus Struve et al. (2009)	intercept	−63.19	−86.77	−39.61	1.50E−07	62.87	95.50	2.53E−148	329.62
Overall minus Gray et al. (2021)	intercept	−56.97	−80.64	−33.31	2.37E−06	59.25	94.78	3.05E−115	311.44

* Indicates model with lowest Akaike information criterion (AIC).
Abbreviations: CI = confidence interval; I^2 = describes the percentage of the variability in effect estimates that is due to heterogeneity rather than sampling error; Tau = estimated standard deviation of the true underlying effect sizes across studies in the random-effects model meta-analysis.

Table 5-4. Comparison of Benchmark Dose Estimates for DBP and Fetal Testosterone in Rats

Analysis	BMR	BMD (mg/kg-day)	CI, Lower Bound (mg/kg-day)	CI, Upper Bound (mg/kg-day)	AIC
2017 NASEM analysis using Metafor Version 2.0.0 (as reported in Tables C6-7 and C6-8 of NASEM (2017))					
Linear in dose100	5%	17	14	22	285.72
Linear in dose100	40%	174	143	222	
Linear Quadratic in dose100*	5%	12	8	22	277.00*
Linear Quadratic in dose100*	40%	125	85	205	
Updated analysis using Metafor Version 2.0.0 including new study by Gray et al. (2021)					
Linear in dose100	5%	20	18	24	344.58
Linear in dose100	10%	42	37	49	
Linear in dose100	40%	204	178	240	
Linear Quadratic in dose100*	5%	15	11	21	334.19*
Linear Quadratic in dose100*	10%	30	23	43	
Linear Quadratic in dose100*	40%	154	119	211	
Updated analysis using Metafor Version 4.6.0 including new study by Gray et al. (2021)					
Linear in dose100	5%	20	16	26	354.71
Linear in dose100	10%	41	33	53	
Linear in dose100	40%	199	162	258	
Linear Quadratic in dose100*	5%	14	9	27	343.82*
Linear Quadratic in dose100*	10%	29	20	54	
Linear Quadratic in dose100*	40%	149	101	247	

* Indicates model with lowest Akaike information criterion (AIC).
Abbreviations: BMD = benchmark dose; BMR = benchmark response; CI = confidence interval

5.2 Di(2-ethylhexyl) Phthalate (DEHP)

EPA identified 29 studies of DEHP evaluating testosterone (Table_Apx B-2). Of these studies, 8 met the criteria outlined in Section 2 for inclusion in the updated meta-analysis (Table 5-5). Seven of the eight studies evaluating fetal rat testicular testosterone content and/or *ex vivo* testosterone production were included in the 2017 NASEM meta-analysis. EPA identified new fetal rat testicular testosterone data from one study (Gray et al., 2021), which was included as part of the updated meta-analysis and BMD modeling analysis for DBP. Table 5-5 provides an overview of the eight studies included in the updated meta-analysis.

Twenty-one studies did not meet the inclusion criteria outlined in Section 2 and were excluded from EPA's updated meta-analysis for various reasons, as outlined in Table_Apx B-2. Of the 21 excluded studies, 7 were excluded by NASEM in 2017 or EPA due to data reporting issues (*e.g.*, N reported as range, not exact value or variance type (SEM, SD) not reported) (Spade et al., 2018; Do et al., 2012; Klinefelter et al., 2012; Vo et al., 2009a; Borch et al., 2006b; Borch et al., 2004; Wilson et al., 2004). Ten studies were excluded, as DEHP was administered outside of the critical window of development, testosterone was measured in a postnatal (not fetal) lifestage, and/or serum (not testis) testosterone was evaluated (Rajagopal et al., 2019; Guo et al., 2013; Li et al., 2012; Gray et al., 2009; Lin et al., 2009; Vo et al., 2009a; Ge et al., 2007; Andrade et al., 2006; Akingbemi et al., 2004; Akingbemi et al., 2001). Two studies were excluded because they evaluated testosterone in mice, not rats (Barakat et al., 2018; Gaido et al., 2007), and the remaining two studies were excluded because they evaluated serum (not testis) testosterone following inhalation (not oral) exposures outside the critical window of development in postnatal (not fetal) rats (Ma et al., 2006; Kurahashi et al., 2005).

For the eight included studies, EPA conducted the updated meta-analysis using random effects models, as implemented in the R Metafor package. Metafor versions 2.0.0 and 4.6.0 were used so that results could be compared. Additionally, the updated analysis included a sensitivity analysis to determine if the meta-analysis was sensitive to leaving out results from individual studies.

Table 5-5. Summary of Studies Included in EPA's Meta-analysis and BMD Modeling Analysis for DEHP

Reference (TSCA Study Quality Rating)	Included in NASEM Meta-analysis and BMD Modeling Analysis?	Brief Study Description	Measured Outcome
(Lin et al., 2008) (Medium)	Yes	Pregnant Long-Evans rats (6–9 dams/group) gavaged with 0, 10, 100, 750 mg/kg-day DEHP on GD 2–20	Fetal testis testosterone content on GD 21
(Martino-Andrade et al., 2008) (Medium)	Yes	Pregnant Wistar rats (7 dams/group) gavaged with 0, 150 mg/kg-day DEHP on GD 13–21	Fetal testis testosterone content on GD 21
(Hannas et al., 2011) (Medium)	Yes	Pregnant Wistar rats (3–6 dams/group) gavaged with 0, 100, 300, 500, 625, 750, 875 mg/kg-day DEHP on GD 14–18	*Ex vivo* fetal testicular testosterone production (3–hour incubation) on GD 18
	Yes	Pregnant SD rats (3–6 dams/group) gavaged with 0, 100, 300, 500, 625, 750, 875 mg/kg-day DEHP on GD 14–18	

Reference (TSCA Study Quality Rating)	Included in NASEM Meta-analysis and BMD Modeling Analysis?	Brief Study Description	Measured Outcome
(Culty et al., 2008) (Medium)	Yes	Pregnant SD rats (3 dams/group) gavaged with 0, 117, 234, 469, 938 mg/kg-day DEHP on GD 14–20	*Ex vivo* fetal testicular testosterone production (24–hour incubation) on GD 21
(Furr et al., 2014) (High)	Yes	Pregnant SD rats (2–3 dams/group) gavaged with 0, 100, 300, 600, 900 mg/kg-day DEHP on GD 14–18 (Block 31)	*Ex vivo* fetal testicular testosterone production (3-hour incubation) on GD 18
	Yes	Pregnant SD rats (2–3 dams/group) gavaged with 0, 100, 300, 600, 900 mg/kg-day DEHP on GD 14–18 (Block 32)	
(Howdeshell et al., 2008) (High)	Yes	Pregnant SD rats (4 dams/group) gavaged with 0, 100, 300, 600, 900 mg/kg-day DEHP on GD 14-18	*Ex vivo* fetal testicular testosterone production (3–hour incubation) on GD 18
(Saillenfait et al., 2013) (High)	Yes	Pregnant SD rats (8–16 dams/group) gavaged with 0, 50, 625 mg/kg-day DEHP on GD 12–19	*Ex vivo* fetal testicular testosterone production (3–hour incubation) on GD 19
(Gray et al., 2021) (High)	No (new study)	Pregnant SD rats (2–3 dams/group) gavaged with 0, 100, 300, 600, 900 mg/kg-day DEHP on GD 14–18 (Block 76).	*Ex vivo* fetal testicular testosterone production (3–hour incubation) on GD 18
	No (new study)	Pregnant SD rats (3 dams/group) gavaged with 0, 100, 300, 600, 900 mg/kg-day DEHP on GD 14–18 (Block 77).	

Overall meta-analyses and sensitivity analyses results obtained using Metafor Versions 2.0.0 and 4.6.0 are shown in Table 5-6 and Table 5-7, respectively. Comparisons of BMD estimates obtained by NASEM (2017) and as part of EPA's updated analysis including new data are shown in Table 5-8. Additional meta-analysis results (*i.e.*, forest plots) and BMD model fit curves are shown in Appendix A.3. For meta-analyses conducted using both versions of Metafor, there was a statistically significant overall effect and linear trends in $\log_{10}$(dose) and dose, with an overall effect that is large in magnitude (>50% change). For both meta-analyses, there was substantial, statistically significant heterogeneity in all cases ($I^2 > 90\%$ for Metafor v.2.0.0; $I^2 > 90\%$ for Metafor v.4.6.0). The statistical significance of these effects was robust to leaving out individual studies for analyses conducted with both versions of Metafor. Although there was substantial heterogeneity, standard deviation of the random effect (tau) was less than the estimated size of the effect at higher doses. Therefore, the heterogeneity does not alter the conclusion that gestational exposure to DEHP reduces fetal testicular testosterone in the rat.

For meta-analyses conducted using both versions of Metafor, the linear-quadratic model provided the best fit (*i.e.*, had lower AIC than the linear model) (Table 5-8). BMD estimates from the linear-quadratic model were 17 mg/kg-day (95% CI: 12, 26) for a 5 percent change (BMR = 5%), 35 mg/kg-day (95% CI: 26, 52) for a 10 percent change (BMR = 10%), and 178 mg/kg-day (95% CI: 134, 251) for a 40 percent change (BMR = 40%) when Metafor Version 2.0.0 was used. Similarly, BMD estimates from

the linear-quadratic model were 17 mg/kg-day (95% CI: 11, 31) for a 5 percent change (BMR = 5%), 35 mg/kg-day (95% CI: 24, 63) for a 10 percent change (BMR = 10%), and 178 mg/kg-day (95% CI: 122, 284) for a 40 percent change (BMR = 40%) when Metafor Version 4.6.0 was used.

Notably, Metafor versions 2.0.0 and 4.6.0 provided identical BMD_5 (17 mg/kg-day), BMD_{10} (35 mg/kg-day), and BMD_{40} (178 mg/kg-day) estimates for the best fitting, linear-quadratic model for the updated analysis including the new data (Table 5-8), and these results are similar to those obtained in the 2017 NASEM meta-analysis (*i.e.*, BMD_5 and BMD_{40} estimates of 15 and 161 mg/kg-day, respectively, based on the best fitting linear quadratic model). At the evaluated BMRs of 5 and 40 percent, inclusion of the new data results in slightly higher BMD_5 and BMD_{40} estimates with similar 95 percent confidence intervals compared to results obtained in the 2017 NASEM analysis.

Table 5-6. Updated Overall Meta-analyses and Sensitivity Analyses of Rat Studies of DEHP and Fetal Testosterone (Metafor Version 2.0.0)

Analysis	Estimate	Beta	CI, Lower Bound	CI, Upper Bound	P value	Tau	I^2	P Value for Heterogeneity	AIC
Primary analysis									
Overall	intercept	−103.69	−127.11	−80.27	4.04E−18	75.18	98.65	5.73E−270	477.69
Trend in log10(dose)	log10(dose)	−135.61	−170.18	−101.03	1.51E−14	46.35	96.47	2.53E−177	432.47
Linear in dose100	dose100	−21.83	−24.55	−19.11	9.90E−56	45.36	96.60	1.03E−164	439.18
Linear Quadratic in dose100	dose100	−30.80	−41.57	−20.03	2.06E−08	44.20	95.91	1.14E−151	429.15*
Linear Quadratic in dose100	I(dose100^2)	1.21	−0.20	2.62	9.15E−02	44.20	95.91	1.14E−151	429.15
Sensitivity analysis									
Overall minus Lin et al. (2008)	intercept	−108.89	−132.57	−85.22	1.95E−19	73.35	98.67	3.02E−264	441.10
Overall minus Saillenfait et al. (2013)	intercept	−103.49	−127.52	−79.45	3.21E−17	75.21	98.61	4.86E−234	454.76
Overall minus Furr et al. (2014)	intercept	−89.06	−112.06	−66.07	3.20E−14	66.18	98.48	3.72E−220	377.11
Overall minus Gray et al. (2021)	intercept	−110.14	−136.73	−83.54	4.76E−16	76.76	98.49	1.55E−166	386.87
Overall minus Hannas et al. (2011)	intercept	−106.48	−136.42	−76.55	3.13E−12	81.07	97.77	1.03E−181	343.54
Overall minus Howdeshell et al. (2008)	intercept	−106.36	−131.60	−81.12	1.47E−16	77.33	98.83	6.46E−270	433.45
Overall minus Culty et al. (2008)	intercept	−99.32	−124.00	−74.65	3.02E−15	75.33	98.75	1.25E−251	431.74
Overall minus Martino-Andrade et al. (2008)	intercept	−105.35	−129.11	−81.59	3.64E−18	75.39	98.68	4.27E−270	466.34

* Indicates model with lowest Akaike information criterion (AIC).
Abbreviations: CI = confidence interval; I^2 = describes the percentage of the variability in effect estimates that is due to heterogeneity rather than sampling error; Tau = estimated standard deviation of the true underlying effect sizes across studies in the random-effects model meta-analysis

Table 5-7. Updated Overall Meta-analyses and Sensitivity Analyses of Rat Studies of DEHP and Fetal Testosterone (Metafor Version 4.6.0)

Analysis	Estimate	Beta	CI, Lower Bound	CI, Upper Bound	P value	Tau	I^2	P Value for Heterogeneity	AIC
Primary analysis									
Overall	intercept	−103.69	−127.11	−80.27	4.04E−18	75.18	98.65	5.73E−270	477.69
Trend in log10(dose)	log10(dose)	−135.61	−170.18	−101.03	1.51E−14	46.35	96.47	2.53E−177	432.47
Linear in dose100	dose100	−21.92	−25.82	−18.02	3.46E−28	67.96	98.46	0.00E00[a]	448.00
Linear Quadratic in dose100	dose100	−30.88	−45.45	−16.31	3.26E−05	61.77	97.86	4.22E−238	435.16*
Linear Quadratic in dose100	I(dose100^2)	1.21	−0.69	3.10	2.13E−01	61.77	97.86	4.22E−238	435.16
Sensitivity analysis									
Overall minus Lin et al. (2008)	intercept	−108.89	−132.57	−85.22	1.95E−19	73.35	98.67	3.02E−264	441.10
Overall minus Saillenfait et al. (2013)	intercept	−103.49	−127.52	−79.45	3.21E−17	75.21	98.61	4.86E−234	454.76
Overall minus Furr et al. (2014)	intercept	−89.06	−112.06	−66.07	3.20E−14	66.18	98.48	3.72E−220	377.11
Overall minus Gray et al. (2021)	intercept	−110.14	−136.73	−83.54	4.76E−16	76.76	98.49	1.55E−166	386.87
Overall minus Hannas et al. (2011)	intercept	−106.48	−136.42	−76.55	3.13E−12	81.07	97.77	1.03E−181	343.54
Overall minus Howdeshell et al. (2008)	intercept	−106.36	−131.60	−81.12	1.47E−16	77.33	98.83	6.46E−270	433.45
Overall minus Culty et al. (2008)	intercept	−99.32	−124.00	−74.65	3.02E−15	75.33	98.75	1.25E−251	431.74
Overall minus Martino-Andrade et al. (2008)	intercept	−105.35	−129.11	−81.59	3.64E−18	75.39	98.68	4.27E−270	466.34

* Indicates model with lowest Akaike information criterion (AIC).
Abbreviations: CI = confidence interval; I^2 = describes the percentage of the variability in effect estimates that is due to heterogeneity rather than sampling error; Tau = estimated standard deviation of the true underlying effect sizes across studies in the random-effects model meta-analysis
[a] p-value too small to calculate and rounded to zero.

Table 5-8. Comparison of Benchmark Dose Estimates for DEHP and Fetal Testosterone in Rats

Analysis	BMR	BMD (mg/kg-day)	CI, Lower Bound (mg/kg-day)	CI, Upper Bound (mg/kg-day)	AIC
2017 NASEM Analysis for all strains of rats using Metafor Version 2.0.0 (as reported in Tables C5-7, C5-8, and C5-9 of NASEM (2017))					
Linear in dose100	5%	22	20	26	358.32
Linear in dose100	40%	222	195	258	
Linear Quadratic in dose100*	5%	15	11	24	348.01*
Linear Quadratic in dose100*	40%	161	118	236	
Updated analysis using Metafor Version 2.0.0 including new study by Gray et al. (2021)					
Linear in dose100	5%	24	21	27	439.18
Linear in dose100	10%	48	43	55	
Linear in dose100	40%	234	208	267	
Linear Quadratic in dose100*	5%	17	12	26	429.15*
Linear Quadratic in dose100*	10%	35	26	52	
Linear Quadratic in dose100*	40%	178	134	251	
Updated analysis using Metafor Version 4.6.0 including new study by Gray et al. (2021)					
Linear in dose100	5%	23	20	28	448.00
Linear in dose100	10%	48	41	58	
Linear in dose100	40%	233	198	283	
Linear Quadratic in dose100*	5%	17	11	31	435.16*
Linear Quadratic in dose100*	10%	35	24	63	
Linear Quadratic in dose100*	40%	178	122	284	

* Indicates model with lowest Akaike information criterion (AIC).
Abbreviations: BMD = benchmark dose; BMR = benchmark response; CI = confidence interval

5.3 Diisobutyl Phthalate (DIBP)[2]

EPA identified seven studies of DIBP evaluating testosterone (Table_Apx B-3). Of these studies, three met the criteria outlined in Section 2 for inclusion in the updated meta-analysis (Table 5-5). Two of the seven studies evaluating fetal rat testicular testosterone content and/or *ex vivo* testosterone production were included in the 2017 NASEM meta-analysis. EPA identified new fetal rat testicular testosterone data from one study (Gray et al., 2021), which was included as part of the updated meta-analysis and BMD modeling analysis for DBP. Table 5-9 provides an overview of the eight studies included in the updated meta-analysis.

Four studies did not meet the inclusion criteria outlined in Section 2 and were excluded from EPA's updated meta-analysis for various reasons, as outlined in Table_Apx B-3. Of the four excluded studies, two were excluded due to data reporting issues (*i.e.*, N reported as range (not exact value) and/or data reported graphically only) (Saillenfait et al., 2017; Borch et al., 2006a) and two were excluded because serum (not testis) testosterone was measured in mice (not rats) (Pan et al., 2017; Wang et al., 2017).

EPA conducted the updated meta-analysis using random effects models, as implemented in the R metafor package. Metafor versions 2.0.0 and 4.6.0 were used so that results could be compared. Additionally, the updated analysis included a sensitivity analysis to determine if the meta-analysis was sensitive to leaving out results from individual studies. In 2017, NASEM did not conduct a sensitivity analysis because there were too few studies available to do so.

Table 5-9. Summary of Studies Included in EPA's Meta-analysis and BMD Modeling Analysis for DIBP

Reference (TSCA Study Quality Rating)	Included in NASEM Meta-analysis and BMD Modeling Analysis?	Brief Study Description	Measured Outcome
(Hannas et al., 2011) (Medium)	Yes	Pregnant SD rats (3 dams/group) gavaged with 0, 100, 300, 600, 900 mg/kg-day DIBP on GD 14–18.	*Ex vivo* fetal testicular testosterone production (3-hour incubation) on GD 18
(Howdeshell et al., 2008) (High)	Yes	Pregnant SD rats (2-8 dams/group) gavaged with 0, 100, 300, 600, 900 mg/kg-day DIBP on GD 8-18.	*Ex vivo* fetal testicular testosterone production (3-hour incubation) on GD 18
(Gray et al., 2021) (High)	No (new study)	Pregnant SD rats (2-3 dams/group) gavaged with 0, 100, 300, 600, 900 mg/kg-day DIBP on GD 14-18 (Block 67 rats).	*Ex vivo* fetal testicular testosterone production (3-hour incubation) on GD 18

Overall meta-analyses and sensitivity analyses results obtained using Metafor Versions 2.0.0 and 4.6.0 are shown in Table 5-10 and Table 5-11, respectively. A comparison of BMD estimates obtained by

[2] In addition to the meta-analysis, EPA also conducted additional BMD modeling of the three individual studies of DIBP reporting reduced fetal testicular testosterone using all standard continuous models in EPA's BMD software (BMDS 3.3.2) (Gray et al., 2021; Hannas et al., 2011; Howdeshell et al., 2008). BMD model results are reported in EPA's *Non-Cancer Human Health Hazard Assessment for Diisobutyl Phthalate (DIBP)* (U.S. EPA, 2025d).

NASEM (2017) and as part of EPA's updated analysis are shown in Table 5-12. Additional meta-analysis results (*i.e.,* forest plots) and BMD model fit curves are shown in Appendix A.4. For meta-analyses conducted using both versions of Metafor, there was a statistically significant overall effect and linear trends in $\log_{10}$(dose) and dose, with an overall effect that is large in magnitude (>50% change). For both meta-analyses, there was substantial, statistically significant heterogeneity in all cases (I^2>50% for Metafor v.2.0.0; I^2 > 65% for Metafor v.4.6.0). The statistical significance of these effects was robust to leaving out individual studies for analyses conducted with both versions of Metafor. Although there was substantial heterogeneity, standard deviation of the random effect (tau) was less than the estimated size of the effect at higher doses. Therefore, the heterogeneity does not alter the conclusion that gestational exposure to DIBP reduces fetal testicular testosterone in the rat.

For meta-analyses conducted using both versions of Metafor, the linear-quadratic model provided the best fit (*i.e.,* had lower AIC than the linear model) (Table 5-12). BMD estimates from the linear-quadratic model were 36 mg/kg-day (95% CI: 23, 79) for a 5 percent change (BMR = 5%), 74 mg/kg-day (95% CI: 47, 140) for a 10 percent change (BMR = 10%), and 326 mg/kg-day (95% CI: 239, 428) for a 40 percent change (BMR = 40%) when Metafor Version 2.0.0 was used. Similarly, BMD estimates were 55 mg/kg-day (95% CI: NA, 266) for a 10 percent change (BMR = 10%) and 270 mg/kg-day (95% CI: 136, 517) for a 40 percent change (BMR = 40%) when Metafor Version 4.6.0 was used. No BMD value could be estimated for a 5 percent change (BMR = 5%), nor could the 95 percent lower confidence limit be estimated for a 10 percent change ($BMDL_{10}$) using Metafor Version 4.6.0. Given that there were only two studies included in the NASEM meta-analysis in 2017, the updated analysis with the addition of the new study by Gray et al. (2021) resulted in a higher BMD and wider confidence interval at both BMRs compared to the NASEM analysis that did not include the new study, although the $BMDL_5$ of 23 mg/kg-day was identical between NASEM's analysis and the updated analysis including the new study, when using Metafor Version 2.0.0.

Table 5-10. Updated Overall Analyses and Sensitivity Analyses of Rat Studies of DIBP and Fetal Testosterone (Metafor Version 2.0.0)

Analysis	Estimate	Beta	CI, Lower Bound	CI, Upper Bound	P value	Tau	I^2	P Value for Heterogeneity	AIC
Primary analysis									
Overall	intercept	−82.21	−122.85	−41.56	7.36E−05	68.02	96.52	4.18E−54	130.45
Trend in log10(dose)	log10(dose)	−165.55	−205.47	−125.64	4.31E−16	19.89	65.48	3.53E−03	106.31
Linear in dose100	dose100	−18.15	−20.60	−15.70	1.09E−47	13.49	60.77	3.93E−03	108.69
Linear Quadratic in dose100	dose100	−13.89	−22.51	−5.28	1.57E−03	11.98	50.83	2.01E−02	104.31*
Linear Quadratic in dose100	I(dose100^2)	−0.55	−1.64	0.54	3.22E−01	11.98	50.83	2.01E−02	104.31
Sensitivity analysis									
Overall minus Gray et al. (2021)	intercept	−82.31	−135.11	−29.52	2.24E−03	71.76	96.96	3.48E−30	87.28
Overall minus Hannas et al. (2011)	intercept	−69.98	−110.63	−29.34	7.39E−04	55.43	95.94	7.26E−37	83.66
Overall minus Howdeshell et al. (2008)	intercept	−94.90	−151.74	−38.06	1.07E−03	78.38	94.86	3.49E−32	88.36

* Indicates model with lowest Akaike information criterion (AIC).
Abbreviations: CI = confidence interval; I^2 = describes the percentage of the variability in effect estimates that is due to heterogeneity rather than sampling error; Tau = estimated standard deviation of the true underlying effect sizes across studies in the random-effects model meta-analysis

Table 5-11. Updated Overall Analyses and Sensitivity Analyses of Rat Studies of DIBP and Fetal Testosterone (Metafor Version 4.6.0)

Analysis	Estimate	Beta	CI, Lower Bound	CI, Upper Bound	P value	Tau	I^2	P value for Heterogeneity	AIC
Primary analysis									
Overall	intercept	−82.21	−122.85	−41.56	7.36E−05	68.02	96.52	4.18E−54	130.45
Trend in log10(dose)	log10(dose)	-165.55	-205.47	-125.64	4.31E−16	19.89	65.48	3.53E−03	106.31
Linear in dose100	dose100	-18.48	-25.14	-11.81	5.50E−08	60.86	96.92	1.55E−111	120.04
Linear Quadratic in dose100	dose100	-19.18	-41.21	2.85	8.79E-02	48.79	94.49	3.45E−39	111.51*
Linear Quadratic in dose100	I(dose100^2)	0.09	−2.70	2.88	9.50E−01	48.79	94.49	3.45E−39	111.51
Sensitivity analysis									
Overall minus Gray et al. (2021)	intercept	−82.31	−135.11	−29.52	2.24E−03	71.76	96.96	3.48E−30	87.28
Overall minus Hannas et al. (2011)	intercept	−69.98	−110.63	−29.34	7.39E−04	55.43	95.94	7.26E−37	83.66
Overall minus Howdeshell et al. (2008)	intercept	−94.90	−151.74	−38.06	1.07E−03	78.38	94.86	3.49E−32	88.36

* Indicates model with lowest Akaike information criterion (AIC).
Abbreviations: CI = confidence interval; I^2 = describes the percentage of the variability in effect estimates that is due to heterogeneity rather than sampling error; Tau = estimated standard deviation of the true underlying effect sizes across studies in the random-effects model meta-analysis

Table 5-12. Comparison of Benchmark Dose Estimates for DIBP and Fetal Testosterone in Rats

Analysis	BMR	BMD (mg/kg-day)	CI, Lower Bound (mg/kg-day)	CI, Upper Bound (mg/kg-day)	AIC
2017 NASEM analysis using Metafor Version 2.0.0 (as reported in Tables C6-11 and C6-12 of NASEM (2017)) [a]					
Linear in dose100*	5%	27	23	34	75.51*
Linear in dose100*	40%	271	225	342	
Linear Quadratic in dose100	5%	43	23	127	77.04
Linear Quadratic in dose100	40%	341	239	453	
Updated analysis using Metafor Version 2.0.0 including new study by (Gray et al., 2021)					
Linear in dose100	5%	28	25	33	108.69
Linear in dose100	10%	58	51	67	
Linear in dose100	40%	281	248	325	
Linear Quadratic in dose100*	5%	36	23	79	104.31*
Linear Quadratic in dose100*	10%	74	47	140	
Linear Quadratic in dose100*	40%	326	239	428	
Updated analysis using Metafor Version 4.6.0 including new study by (Gray et al., 2021)					
Linear in dose100	5%	28	20	43	120.04
Linear in dose100	10%	57	42	89	
Linear in dose100	40%	276	203	432	
Linear Quadratic in dose100*	5%	NA [c]	NA [b]	207	111.51*
Linear Quadratic in dose100*	10%	55	NA [b]	266	
Linear Quadratic in dose100*	40%	270	136	517	

* Indicates model with lowest Akaike information criterion (AIC).

Abbreviations: BMD = benchmark dose; BMR = benchmark response; CI = confidence interval

[a] EPA noted an apparent discrepancy in the NASEM (2017) report. In Table 3-26, NASEM (2017) notes that no BMD/BMDL estimates could be generated at the 5% response level for DIBP because "the 5% change was well below the range of the data, but it will be 10 times lower because a linear model was used." However, in Table C6-12 of the NASEM (2017) report, BMD/BMDL estimates at the 5% response level are provided for DIBP for the best-fit linear model. In EPA's replicate analysis, identical BMD/BMDL estimates for the 5% response level were obtained. Therefore, BMD/BMDL estimates at the 5% response level for DIBP are reported in this table.

Analysis	BMR	BMD (mg/kg-day)	CI, Lower Bound (mg/kg-day)	CI, Upper Bound (mg/kg-day)	AIC
[b] Estimate could not be derived.					

5.4 Butyl Benzyl Phthalate (BBP)[3]

EPA identified nine studies of BBP evaluating testosterone (Table_Apx B-4). Of these studies, three met the criteria outlined in Section 2 for inclusion in the updated meta-analysis (Table 5-13). Two of the three studies evaluating fetal rat testicular testosterone content and/or *ex vivo* testosterone production were included in the 2017 NASEM meta-analysis (Furr et al., 2014; Howdeshell et al., 2008). EPA identified new fetal rat testicular testosterone data from one study (Gray et al., 2021), which was included as part of the updated meta-analysis and BMD modeling analysis for BBP. Table 5-13 provides an overview of the three studies included in the updated meta-analysis.

Six studies did not meet the inclusion criteria outlined in Section 2 and were excluded from EPA's updated meta-analysis for various reasons, as outlined in Table_Apx B-4. Of the six excluded studies, three were excluded because they measured serum (not testis) testosterone in postnatal (not fetal) lifestages (Ahmad et al., 2014; Aso et al., 2005; Nagao et al., 2000), two were excluded due to data reporting issues (Spade et al., 2018; Wilson et al., 2004), and one was excluded because serum (not testis) testosterone was evaluated in postnatal mice (not fetal rats) (Schmitt et al., 2016).

EPA conducted the updated meta-analysis using random effects models, as implemented in the R Metafor package. Metafor versions 2.0.0 and 4.6.0 were used so that results could be compared. Additionally, the updated analysis included a sensitivity analysis to determine if the meta-analysis was sensitive to leaving out results from individual studies. In 2017, NASEM did not conduct a sensitivity analysis because there were too few studies available to do so.

Table 5-13. Summary of Studies Included in EPA's Meta-analysis and BMD Modeling Analysis for BBP

Reference (TSCA Study Quality Rating)	Included in NASEM Meta-analysis and BMD Modeling Analysis?	Brief Study Description	Measured Outcome
(Howdeshell et al., 2008) (High)	Yes	Pregnant SD rats (2–9 dams/group) gavaged with 0, 100, 300, 600, 900 mg/kg-day BBP on GD 8–18.	*Ex vivo* fetal testicular testosterone production (2–hour incubation) on GD 18
(Furr et al., 2014) (High)	Yes	Pregnant SD rats (2–3 dams/group) gavaged with 0, 100, 300, 600, 900 mg/kg-day BBP on GD 14–18 (Block 36 rats).	*Ex vivo* fetal testicular testosterone production (3–hour incubation) on GD 18
	Yes	Pregnant SD rats (3–4 dams/group) gavaged with 0, 11, 33, 100 mg/kg-day BBP on GD 14–18 (Block 37 rats).	*Ex vivo* fetal testicular testosterone production (3–hour incubation) on GD 18
(Gray et al., 2021) (High)	No (new study)	Pregnant SD rats (3 dams/group) gavaged with 0, 100, 300, 600, 900 mg/kg-day BBP on GD 14–18 (Block 78 rats).	*Ex vivo* fetal testicular testosterone production (3–hour incubation) on GD 18

[3] In addition to the meta-analysis, EPA also conducted additional BMD modeling of the four individual studies of BBP reporting reduced fetal testicular testosterone using all standard continuous models in EPA's BMD software (BMDS 3.3.2) (Gray et al., 2021; Furr et al., 2014; Howdeshell et al., 2008). BMD model results are reported in EPA's *Non-Cancer Human Health Hazard Assessment for Butyl Benzyl Phthalate (BBP)* (U.S. EPA, 2025a).

Overall meta-analyses and sensitivity analyses results obtained using Metafor Versions 2.0.0 and 4.6.0 are shown in Table 5-14 and Table 5-15, respectively. A comparison of BMD estimates obtained by NASEM (2017) and as part of EPA's updated analysis are shown in Table 5-16. Additional meta-analysis results (*i.e.,* forest plots) and BMD model fit curves are shown in Appendix A.5. For meta-analyses conducted using both versions of Metafor, there was a statistically significant overall effect and linear trends in $\log_{10}$(dose) and dose, with an overall effect that is large in magnitude (>50% change). For both meta-analyses, there was substantial, statistically significant heterogeneity in all cases (I^2 > 50% for Metafor v.2.0.0; I^2 > 90% for Metafor v.4.6.0). The statistical significance of these effects was robust to leaving out individual studies for analyses conducted with both versions of Metafor. Although there was substantial heterogeneity, standard deviation of the random effect (tau) was less than the estimated size of the effect at higher doses. Therefore, the heterogeneity does not alter the conclusion that gestational exposure to BBP reduces fetal testicular testosterone in the rat.

For meta-analyses conducted using both versions of Metafor, the linear-quadratic model provided the best fit (*i.e.,* had lower AIC than the linear model) (Table 5-16). BMD estimates from the linear-quadratic model were 31 mg/kg-day (95% CI: 17, 103) for a 5 percent change (BMR = 5%), 63 mg/kg-day (95% CI: 36, 163) for a 10 percent change (BMR = 10%), and 276 mg/kg-day (95% CI: 179, 408) for a 40 percent change (BMR = 40%) when Metafor Version 2.0.0 was used. Similarly, a BMD of 284 mg/kg-day (95% CI: 150, 481) for a 40 percent change (BMR = 40%) was estimated using Metafor Version 4.6.0; however, no BMD estimates could be derived for 5 and 10 percent changes (BMRs = 5 and 10%) using Metafor Version 4.6.0. Again, inclusion of the new study by Gray et al. (2021) resulted in a higher BMD at both response rates, although the $BMDL_5$ for EPA's updated analysis including the new study (17 mg/kg-day) was similar to the NASEM 2017 analysis when both are compared using Metafor Version 2.0.0 (13 mg/kg-day).

Table 5-14. Updated Overall Meta-analyses and Sensitivity Analyses of Rat Studies of BBP and Fetal Testosterone (Metafor Version 2.0.0)

Analysis	Estimate	Beta	CI, Lower Bound	CI, Upper Bound	P Value	Tau	I^2	P Value for Heterogeneity	AIC
Primary analysis									
Overall	intercept	−83.62	−127.17	−40.06	1.68E−04	83.98	98.20	4.78E−151	169.89
Trend in log10(dose)	log10(dose)	−120.36	−169.45	−71.28	1.54E−06	49.93	94.66	3.34E−36	149.12
Linear in dose100	dose100	−22.64	−26.33	−18.96	2.10E−33	29.83	86.32	2.75E−22	143.19
Linear Quadratic in dose100	dose100	−16.12	−29.93	−2.30	2.22E−02	30.72	84.75	1.74E−20	136.90*
Linear Quadratic in dose100	I(dose100^2)	−0.87	−2.64	0.90	3.35E−01	30.72	84.75	1.74E−20	136.90
Sensitivity analysis									
Overall minus Furr et al. (2014)	intercept	−90.83	−160.08	−21.59	1.01E−02	97.63	97.87	2.72E−33	91.46
Overall minus Gray et al. (2021)	intercept	−78.47	−125.70	−31.24	1.13E−03	77.72	98.17	5.38E−125	122.09
Overall minus Howdeshell et al. (2008)	intercept	−84.05	−134.86	−33.24	1.19E−03	84.27	98.27	8.30E−102	123.25

* Indicates model with lowest Akaike information criterion (AIC).
Abbreviations: CI = confidence interval; I^2 = describes the percentage of the variability in effect estimates that is due to heterogeneity rather than sampling error; Tau = estimated standard deviation of the true underlying effect sizes across studies in the random-effects model meta-analysis

Table 5-15. Updated Overall Meta-analyses and Sensitivity Analyses of Rat Studies of BBP and Fetal Testosterone (Metafor Version 4.6.0)

Analysis	Estimate	Beta	CI, Lower Bound	CI, Upper Bound	P Value	Tau	I^2	P value for Heterogeneity	AIC
Primary analysis									
Overall	intercept	−83.62	−127.17	−40.06	1.68E−04	83.98	98.20	4.78E−151	169.89
Trend in log10(dose)	log10(dose)	−120.36	−169.45	−71.28	1.54E−06	49.93	94.66	3.34E−36	149.12
Linear in dose100	dose100	−22.98	−30.32	−15.63	8.69E−10	69.12	97.13	7.81E−82	153.33
Linear Quadratic in dose100	dose100	−15.00	−36.40	6.40	1.70E−01	50.89	93.85	8.24E−53	140.94*
Linear Quadratic in dose100	I(dose100^2)	−1.04	−3.78	1.69	4.54E−01	50.89	93.85	8.24E−53	140.94
Sensitivity analysis									
Overall minus Furr et al. (2014)	intercept	−90.83	−160.08	−21.59	1.01E−02	97.63	97.87	2.72E−33	91.46
Overall minus Gray et al. (2021)	intercept	−78.47	−125.70	−31.24	1.13E−03	77.72	98.17	5.38E−125	122.09
Overall minus Howdeshell et al. (2008)	intercept	−84.05	−134.86	−33.24	1.19E−03	84.27	98.27	8.30E−102	123.25

* Indicates model with lowest Akaike information criterion (AIC).
Abbreviations: CI = confidence interval; I^2 = describes the percentage of the variability in effect estimates that is due to heterogeneity rather than sampling error; Tau = estimated standard deviation of the true underlying effect sizes across studies in the random-effects model meta-analysis

Table 5-16. Comparison of Benchmark Dose Estimates for BBP and Fetal Testosterone in Rats

Analysis	BMR	BMD (mg/kg-day)	CI, Lower Bound (mg/kg-day)	CI, Upper Bound (mg/kg-day)	AIC
2017 NASEM analysis using Metafor Version 2.0.0 (as reported in Tables C6-3 and C6-4 of NASEM, (2017))					
Linear in dose100	5%	23	19	29	103.86
Linear in dose100	40%	231	192	290	
Linear Quadratic in dose100*	5%	23	13	74	100.00*
Linear Quadratic in dose100*	40%	228	140	389	
Updated analysis using Metafor Version 2.0.0 including new study by (2021)					
Linear in dose100	5%	23	19	27	143.19
Linear in dose100	10%	47	40	56	
Linear in dose100	40%	226	194	269	
Linear Quadratic in dose100*	5%	31	17	103	136.90*
Linear Quadratic in dose100*	10%	63	36	163	
Linear Quadratic in dose100*	40%	276	179	408	
Updated analysis using Metafor Version 4.6.0 including new study by (2021)					
Linear in dose100	5%	22	17	33	153.33
Linear in dose100	10%	46	35	67	
Linear in dose100	40%	222	168	327	
Linear Quadratic in dose100*	5%	NA [a]	NA [b]	236	140.94*
Linear Quadratic in dose100*	10%	NA [a]	NA [b]	280	
Linear Quadratic in dose100*	40%	284	150	481	

* Indicates model with lowest Akaike information criterion (AIC).
Abbreviations: BMD = benchmark dose; BMR = benchmark response; CI = confidence interval
[a] BMD and BMDL estimates could not be derived.

5.5 Dicyclohexyl Phthalate (DCHP)[4]

NASEM (2017) did not include DCHP as part of its phthalate meta-analysis. EPA identified seven studies of DCHP evaluating testosterone (Table_Apx B-5). Of these studies, two met the criteria outlined in Section 2 for inclusion in the meta-analysis (Gray et al., 2021; Furr et al., 2014) (Table 5-17). Five studies were excluded from the meta-analysis because they evaluated serum (not testis) testosterone and/or testosterone was measured during a postnatal (not fetal) lifestage (Lv et al., 2019; Li et al., 2016; Ahbab and Barlas, 2015, 2013; Hoshino et al., 2005) (Table_Apx B-5). Meta-analyses were conducted using Metafor Versions 2.0.0 and 4.6.0 so that results could be compared. No sensitivity analysis was conducted because too few studies were available to do so.

Table 5-17. Summary of Studies Included in EPA's Meta-analysis and BMD Modeling Analysis for DCHP

Reference (TSCA Study Quality Rating)	Included in NASEM Meta-analysis and BMD Modeling Analysis?	Brief Study Description	Measured Outcome
(Furr et al., 2014) (High)	No	Pregnant SD rats (3–4 dams/group) gavaged with 0, 33, 100, 300 mg/kg-day DCHP on GD 14-18 (Block 33).	*Ex vivo* fetal testicular testosterone production (3–hour incubation) on GD 18
	No	Pregnant SD rats (2–3 dams/group) gavaged with 0, 100, 300, 600, 900 mg/kg-day DCHP on GD 14–18 (Block 23).	*Ex vivo* fetal testicular testosterone production (3–hour incubation) on GD 18
(Gray et al., 2021) (High)	No	Pregnant SD rats (3 dams/group) gavaged with 0, 100, 300, 600, 900 mg/kg-day DCHP on GD 14–18 (Block 148).	*Ex vivo* fetal testicular testosterone production (3–hour incubation) on GD 18

Overall meta-analysis results obtained using Metafor Versions 2.0.0 and 4.6.0 are shown in Table 5-18 and Table 5-19, respectively, while a comparison of BMD estimates obtained using both versions of Metafor are shown in Table 5-20. Additional meta-analysis results (*i.e.*, forest plots) and BMD model fit curves are shown in Appendix A.6. Metafor Versions 2.0.0 and 4.6.0 provided similar meta- analysis and BMD modeling results for DCHP. For meta-analysis conducted using both versions of Metafor, there was a statistically significant overall effect and linear trends in $\log_{10}$(dose) and dose, with an overall effect that is large in magnitude (>50% change). For both meta-analysis, there was substantial, statistically significant heterogeneity in all cases (I^2 > 75% for Metafor v.2.0.0; I^2 > 80% for Metafor v.4.6.0). Although there was substantial heterogeneity, standard deviation of the random effect (tau) was less than the estimated size of the effect at higher doses. Therefore, the heterogeneity does not alter the conclusion that gestational exposure to DCHP reduces fetal testicular testosterone in the rat.

For meta-analyses conducted using both versions of Metafor, the linear-quadratic model provided the best fit (*i.e.*, had lower AIC than the linear model) (Table 5-20). BMD estimates from the linear-quadratic model were 8.2 mg/kg-day (95% CI: 6.5, 11) for a 5 percent change (BMR = 5%), 17 mg/kg-

[4] In addition to the meta-analysis, EPA also conducted additional BMD modeling of all individual studies of DCHP in Table 5-17 reporting reduced fetal testicular testosterone using all standard continuous models in EPA's BMD software (BMDS Online Version 25.1). These BMD model results are reported in EPA's *Non-Cancer Human Health Hazard Assessment for Dicyclohexyl Phthalate (DCHP)* (U.S. EPA, 2025c).

day (95% CI: 13, 23) for a 10 percent change (BMR = 10%), and 88 mg/kg-day (95% CI: 69, 121) for a 40 percent change (BMR = 40%) when Metafor Version 2.0.0 was used. Similarly, BMD estimates were 8.4 mg/kg-day (95% CI: 6.0, 14) for a 5 percent change (BMR = 5%), 17 mg/kg-day (95% CI: 12, 29) for a 10 percent change (BMR = 10%), and 90 mg/kg-day (95% CI: 63, 151) for a 40 percent change (BMR = 40%) when Metafor Version 4.6.0 was used.

Notably, Metafor versions 2.0.0 and 4.6.0 provided similar BMD_5 (8.2 vs. 8.4 mg/kg-day), BMD_{10} (17 mg/kg-day for both versions of Metafor), and BMD_{40} (88 vs. 90 mg/kg-day) estimates for the best fitting, linear-quadratic model (Table 5-20).

Table 5-18. Overall Meta-analyses of Rat Studies of DCHP and Fetal Testosterone (Metafor Version 2.0.0)

Analysis	Estimate	Beta	CI, Lower Bound	CI, Upper Bound	P Value	Tau	I^2	P Value for Heterogeneity	AIC
Primary analysis									
Overall	intrcpt	−113.99	−146.03	−81.95	3.1E−12	50.13	88.36	3.6E−12	114.46
Trend in log10(dose)	log10(dose)	−77.00	−135.97	−18.04	1.0E−02	39.19	81.97	5.5E−08	104.45
Linear in dose100	dose100	−22.30	−31.07	−13.52	6.4E−07	68.41	93.45	2.3E−32	119.27
Linear Quadratic in dose100	dose100	−62.86	−79.25	−46.47	5.7E−14	32.05	75.41	7.6E−05	103.12*
Linear Quadratic in dose100	I(dose100^2)	5.64	3.48	7.79	2.9E−07	32.05	75.41	7.6E−05	103.12

* Indicates model with lowest Akaike information criterion (AIC).

Abbreviations: CI = confidence interval; I^2 = describes the percentage of the variability in effect estimates that is due to heterogeneity rather than sampling error; Tau = estimated standard deviation of the true underlying effect sizes across studies in the random-effects model meta-analysis

Table 5-19. Overall Meta-analyses of Rat Studies of DCHP and Fetal Testosterone (Metafor Version 4.6.0)

Analysis	Estimate	Beta	CI, Lower Bound	CI, Upper Bound	P value	Tau	I^2	P Value for Heterogeneity	AIC
Overall	intrcpt	−113.99	−146.03	−81.95	3.1E−12	50.13	88.36	3.6E−12	114.46
Trend in log10(dose)	log10(dose)	−77.00	−135.97	−18.04	1.0E−02	39.19	81.97	5.5E−08	104.45
Linear in dose100	dose100	−22.14	−28.75	−15.54	5.0E−11	49.12	88.03	8.1E−13	121.53
Linear Quadratic in dose100	dose100	−61.83	−86.20	−37.46	6.6E−07	51.94	88.95	1.4E−12	104.92*
Linear Quadratic in dose100	I(dose100^2)	5.39	2.21	8.56	8.8E−04	51.94	88.95	1.4E−12	104.92

* Indicates model with lowest Akaike information criterion (AIC).

Abbreviations: CI = confidence interval; I^2 = describes the percentage of the variability in effect estimates that is due to heterogeneity rather than sampling error; Tau = estimated standard deviation of the true underlying effect sizes across studies in the random-effects model meta-analysis

Table 5-20. Comparison of Benchmark Dose Estimates for DCHP and Fetal Testosterone in Rats

Analysis	BMR	BMD (mg/kg-day)	CI, Lower Bound (mg/kg-day)	CI, Upper Bound (mg/kg-day)	AIC
Analysis using Metafor Version 2.0.0					
Linear in dose100	5%	23	17	38	
Linear in dose100	10%	47	34	78	119.27
Linear in dose100	40%	229	164	378	
Linear Quadratic in dose100*	5%	8.2	6.5	11	
Linear Quadratic in dose100*	10%	17	13	23	103.12*
Linear Quadratic in dose100*	40%	88	69	121	
Analysis using Metafor Version 4.6.0					
Linear in dose100	5%	23	18	33	
Linear in dose100	10%	48	37	68	121.53
Linear in dose100	40%	231	178	329	
Linear Quadratic in dose100*	5%	8.4	6.0	14	
Linear Quadratic in dose100*	10%	17	12	29	104.92*
Linear Quadratic in dose100*	40%	90	63	151	

* Indicates model with lowest Akaike information criterion (AIC).
Abbreviations: BMD = benchmark dose; BMR = benchmark response; CI = confidence interval

6 COMPARISON OF BENCHMARK DOSE ESTIMATES

Table 6-1 compares NASEM and EPA's updated BMD modeling results (reported herein) for decreased fetal testicular testosterone in rats for DBP, DEHP, DIBP, BBP, and DCHP. Table 6-1 also includes NASEM and EPA's updated BMD modeling results for DINP, which are reported in EPA's *Non-Cancer Human Health Hazard Assessment for Diisononyl Phthalate (DINP)* (U.S. EPA, 2025e) to allow for a comparison of BMD modeling results for all phthalates for which modeling of fetal testicular testosterone was conducted. As can be seen from Table 6-1 and as discussed further below, EPA's updated meta-analysis and BMD modeling results generated using Metafor Version 2.0.0 and 4.6.0 are similar for DEHP, DBP, DCHP, and DINP at the evaluated BMRs of 5, 10, and 40 percent. In contrast, for BBP and DIBP, Metafor Version 2.0.0 and 4.6.0 provided differing results. The following similarities and differences are apparent based on BMD/BMDL results provided in Table 6-1.

- ***DBP:*** The linear-quadratic model provided the best fit (based on lowest AIC), regardless of which version of Metafor was used. For EPA's updated analysis, BMD/BMDL estimates at the 5, 10, and 40 percent response levels are similar, regardless of which version of Metafor was used. BMD/BMDL estimates at the 5, 10, and 40 percent response levels are: 15/11, 30/23, and 154/119 mg/kg-day, respectively, using Metafor version 2.0.0 compared to 14/9, 29/20, and 149/101 mg/kg-day, respectively, using Metafor version 4.6.0. These results are similar to the BMD/BMDL estimates of 12/8 and 125/85 mg/kg-day at the 5 and 40 percent response levels, respectively, reported by NASEM (2017).
- ***DEHP:*** The linear-quadratic model provided the best fit (based on lowest AIC), regardless of which version of Metafor was used. For EPA's updated analysis, BMD/BMDL estimates at the 5, 10, and 40 percent response levels are similar, regardless of which version of Metafor was used. BMD/BMDL estimates at the 5, 10, and 40 percent response levels are: 17/12, 35/26, and 178/134 mg/kg-day, respectively, using Metafor version 2.0.0 compared to 17/11, 35/24, and 178/122 mg/kg-day, respectively, using Metafor version 4.6.0. These results are similar to the BMD/BMDL estimates of 15/11 and 161/118 mg/kg-day at the 5 and 40 percent response levels, respectively, reported by NASEM (2017).
- ***DIBP:*** For EPA's updated analysis, the linear-quadratic model provided the best fit (based on lowest AIC), regardless of which version of Metafor was used. For EPA's updated analysis, BMD/BMDL estimates differed depending on which version of Metafor was used. BMD/BMDL estimates at the 5, 10, and 40 percent response levels are: 36/23, 74/47, and 326/239 mg/kg-day, respectively using Metafor version 2.0.0. These results are similar to the BMD/BMDL estimates of 27/23 and 271/225 mg/kg-day at the 5 and 40 percent response levels, respectively, reported by NASEM (2017), however, in the NASEM (2017) the linear model provide the best fit (based on lowest AIC). When Metafor Version 4.6.0 was used, similar BMD/BMDL results were obtained at the 40 percent response level (BMD_{40}/$BMDL_{40}$ = 279/136 mg/kg-day). At the 10 percent response level, the BMD was estimated to 55 mg/kg-day, however, no $BMDL_{10}$ could be estimated. Similarly, no BMD/BMDL estimates could be generated at the 5 percent response level using Metafor Version 4.6.0. Presently, the exact reason(s) why BMD and/or BMDL estimates could not be generated at the 5 or 10 percent response levels are unclear. As described in Section 3 of this document, many updates have been made to the Metafor Version 4.6.0 since Version 2.0.0.
- ***BBP:*** The linear-quadratic model provided the best fit (based on lowest AIC), regardless of which version of Metafor was used. For EPA's updated analysis, BMD/BMDL estimates differed depending on which version of Metafor was used. BMD/BMDL estimates at the 5, 10, and 40 percent response levels are: 31/17, 63/36, and 276/179 mg/kg-day, respectively using Metafor version 2.0.0. These results are similar to the BMD/BMDL estimates of 23/13 and

228/140 mg/kg-day at the 5 and 40 percent response levels, respectively, reported by NASEM (2017). When Metafor Version 4.6.0 was used, similar BMD/BMDL results were obtained at the 40 percent response level ($BMD_{40}/BMDL_{40}$ = 284/150 mg/kg-day), however, no BMD/BMDL estimates could be generated at the 5 or 10 percent response levels. Presently, the precise reason(s) why BMD/BMDL estimates could not be generated at the 5 or 10 percent response levels are unclear. As described in Section 3 of this document, many updates have been made to the Metafor Version 4.6.0 since Version 2.0.0.

- ***DCHP:*** The linear-quadratic model provided the best fit (based on lowest AIC), regardless of which version of Metafor was used. For EPA's updated analysis, BMD/BMDL estimates at the 5, 10, and 40 percent response levels are similar, regardless of which version of Metafor was used. BMD/BMDL estimates at the 5, 10, and 40 percent response levels are: 8.2/6.5, 17/13, and 88/69 mg/kg-day, respectively, using Metafor version 2.0.0 compared to 8.4/6.0, 17/12, and 90/63 mg/kg-day, respectively, using Metafor version 4.6.0. NASEM (2017) did not include DCHP in its 2017 analysis.
- ***DINP:*** The linear-quadratic model provided the best fit (based on lowest AIC), regardless of which version of Metafor was used. For EPA's updated analysis, BMD/BMDL estimates at the 5, 10, and 40 percent response levels are similar, regardless of which version of Metafor was used. BMD/BMDL estimates at the 5, 10, and 40 percent response levels are: 79/52, 160/108, and 715/584 mg/kg-day, respectively, using Metafor version 2.0.0 compared to 74/47, 152/97, and 699/539 mg/kg-day, respectively, using Metafor version 4.6.0. These results are similar to the BMD/BMDL estimates of 76/49 and 701/552 mg/kg-day at the 5 and 40 percent response levels, respectively, reported by NASEM (2017). (*Note: see EPA's Non-Cancer Human Health Hazard Assessment for Diisononyl Phthalate (DINP) (U.S. EPA, 2025e) for Meta-analysis and BMD Model Results.*)

Table 6-1. Comparison of BMD Modeling Results for DEHP, DBP, DIBP, BBP, DCHP, and DINP

Phthalate	Model Providing Best Fit	NASEM (2017) Analysis (Metafor Version 2.0.0)		EPA Updated Analysis (Metafor Version 2.0.0)			EPA Updated Analysis (Metafor Version 4.6.0)		
		BMD_5 Estimates (mg/kg-day) [95% CI]	BMD_{40} Estimates (mg/kg-day) [95% CI]	BMD_5 Estimates (mg/kg-day) [95% CI]	BMD_{10} Estimates (mg/kg-day) [95% CI]	BMD_{40} Estimates (mg/kg-day) [95% CI]	BMD_5 Estimates (mg/kg-day) [95% CI]	BMD_{10} Estimates (mg/kg-day) [95% CI]	BMD_{40} Estimates (mg/kg-day) [95% CI]
DBP	Linear Quadradic [a]	12 [8, 22]	125 [85, 205]	15 [11, 21]	30 [23, 43]	154 [119, 211]	14 [9, 27]	29 [20, 54]	149 [101, 247]
DEHP	Linear Quadradic [a]	15 [11, 24]	161 [118, 236]	17 [12, 26]	35 [26, 52]	178 [134, 251]	17 [11, 31]	35 [24, 63]	178 [122, 284]
DIBP	Linear Quadradic [a b]	27 [23, 34] [b]	271 [225, 342] [b]	36 [23, 79]	74 [47, 140]	326 [239, 428]	– [c]	55 [NA, 266] [c]	279 [136, 517]
BBP	Linear Quadradic [a]	23 [13, 74]	228 [140, 389]	31 [17, 103]	63 [36, 163]	276 [179, 408]	– [c]	– [c]	284 [150, 481]
DCHP	Linear Quadradic [a]	– [d]	– [d]	8.2 [6.5, 11]	17 [13, 23]	88 [69, 121]	8.4 [6.0, 14]	17 [12, 29]	90 [63, 151]
DINP [e]	Linear Quadradic [a]	76 [49, 145]	701 [552, 847]	79 [52, 145]	160 [108, 262]	715 [584, 842]	74 [47, 158]	152 [97, 278]	699 [539, 858]

Abbreviations: BMD = benchmark dose associated with 5% (BMD_5), 10% (BMD_{10}) or 40% (BMD_{40}) response level; CI = confidence interval

[a] Unless otherwise noted, the linear quadratic model provided the best fit (based on lowest AIC) for NASEM and EPA updated analyses using Metafor versions 2.0.0 and 4.6.0.

[b] Linear model provided the best fit (bast on lowest AIC) for NASEM (2017) modeling of DIBP.

[c] BMD and/or BMDL estimate could not be derived.

[d] DCHP was not included in the 2017 NASEM meta-analysis.

[e] See EPA's *Non-Cancer Human Health Hazard Assessment for Diisononyl Phthalate (DINP)* (U.S. EPA, 2025e) for meta-analysis and BMD model results.

7 CONCLUSION

Herein, EPA conducted an updated meta-analysis and BMD modeling analysis of decreased fetal testicular testosterone in rats. This analysis represents an update of the analysis conducted by NASEM (2017). As part of the updated analysis, EPA conducted modeling using Metafor Version 2.0.0 (version originally used by NASEM in 2017) and Version 4.6.0 (most recent version available at the time of EPA's updated analysis). EPA also evaluated BMRs of 5, 10, and 40 percent. Comparatively, NASEM (2017) evaluated BMRs of 5 and 40 percent. As discussed in Section 6, similar BMD/BMDL estimates at the 5, 10, and 40 percent response levels were obtained using Metafor Version 2.0.0 and 4.6.0 for DEHP, DBP, DCHP, and DINP. However, for DIBP and BBP, Metafor Version 2.0.0 and 4.6.0 provided differing results, particularly at the 5 and 10 percent response levels, where BMD and/or BMDL estimates could not be generated using Metafor Version 4.6.0. The precise reason(s) for the differing results for DIBP and BBP using Metafor Version 2.0.0 and 4.6.0 are unclear. As described in Section 3 of this document, many updates have been made to Metafor Version 4.6.0 since Version 2.0.0.

Overall, EPA selected BMD modeling results obtained using Metafor Version 4.6.0 for use in the single phthalate risk evaluations and phthalate cumulative risk assessment because these results were obtained using the most up-to-date version of the Metafor package available at the time of the updated meta-analysis and BMD modeling analysis.

This TSD was released for public comment and was peer-reviewed by the Science Advisory Committee on Chemicals (SACC) during the August 4 to 8, 2025 SACC Meeting (U.S. EPA, 2025k). Following SACC peer-review and public comment, this technical support document was revised to incorporate recommendations from the SACC and public commenters. Readers are directed to EPA's response to public comments summary document and EPA's response to the 2025 phthalates SACC meeting report for further details.

REFERENCES

Ahbab, MA; Barlas, N. (2013). Developmental effects of prenatal di-n-hexyl phthalate and dicyclohexyl phthalate exposure on reproductive tract of male rats: Postnatal outcomes. Food Chem Toxicol 51: 123-136. https://dx.doi.org/10.1016/j.fct.2012.09.010

Ahbab, MA; Barlas, N. (2015). Influence of in utero di-n-hexyl phthalate and dicyclohexyl phthalate on fetal testicular development in rats. Toxicol Lett 233: 125-137. http://dx.doi.org/10.1016/j.toxlet.2015.01.015

Ahmad, R; Gautam, AK; Verma, Y; Sedha, S; Kumar, S. (2014). Effects of in utero di-butyl phthalate and butyl benzyl phthalate exposure on offspring development and male reproduction of rat. Environ Sci Pollut Res Int 21: 3156-3165. https://dx.doi.org/10.1007/s11356-013-2281-x

Akingbemi, BT; Ge, R; Klinefelter, GR; Zirkin, BR; Hardy, MP. (2004). Phthalate-induced Leydig cell hyperplasia is associated with multiple endocrine disturbances. Proc Natl Acad Sci USA 101: 775-780. http://dx.doi.org/10.1073/pnas.0305977101

Akingbemi, BT; Youker, RT; Sottas, CM; Ge, R; Katz, E; Klinefelter, GR; Zirkin, BR; Hardy, MP. (2001). Modulation of rat Leydig cell steroidogenic function by di(2-ethylhexyl)phthalate. Biol Reprod 65: 1252-1259. http://dx.doi.org/10.1095/biolreprod65.4.1252

Allen, BC; Kavlock, RJ; Kimmel, CA; Faustman, EM. (1994a). Dose-response assessment for developmental toxicity II: Comparison of generic benchmark dose estimates with no observed adverse effect levels. Fundam Appl Toxicol 23: 487-495. https://dx.doi.org/10.1006/faat.1994.1133

Allen, BC; Kavlock, RJ; Kimmel, CA; Faustman, EM. (1994b). Dose-response assessment for developmental toxicity III: statistical models. Fundam Appl Toxicol 23: 496-509. https://dx.doi.org/10.1006/faat.1994.1134

Andrade, AJ; Grande, SW; Talsness, CE; Grote, K; Golombiewski, A; Sterner-Kock, A; Chahoud, I. (2006). A dose-response study following in utero and lactational exposure to di-(2-ethylhexyl) phthalate (DEHP): Effects on androgenic status, developmental landmarks and testicular histology in male offspring rats. Toxicology 225: 64-74. http://dx.doi.org/10.1016/j.tox.2006.05.007

Aso, S; Ehara, H; Miyata, K; Hosyuyama, S; Shiraishi, K; Umano, T; Minobe, Y. (2005). A two-generation reproductive toxicity study of butyl benzyl phthalate in rats. J Toxicol Sci 30: S39-S58. https://dx.doi.org/10.2131/jts.30.S39

Barakat, R; Lin, PC; Park, CJ; Best-Popescu, C; Bakry, HH; Abosalem, ME; Abdelaleem, NM; Flaws, JA; Ko, C. (2018). Prenatal Exposure to DEHP Induces Neuronal Degeneration and Neurobehavioral Abnormalities in Adult Male Mice. Toxicol Sci 164: 439-452. http://dx.doi.org/10.1093/toxsci/kfy103

Borch, J; Axelstad, M; Vinggaard, AM; Dalgaard, M. (2006a). Diisobutyl phthalate has comparable anti-androgenic effects to di-n-butyl phthalate in fetal rat testis. Toxicol Lett 163: 183-190. http://dx.doi.org/10.1016/j.toxlet.2005.10.020

Borch, J; Ladefoged, O; Hass, U; Vinggaard, AM. (2004). Steroidogenesis in fetal male rats is reduced by DEHP and DINP, but endocrine effects of DEHP are not modulated by DEHA in fetal, prepubertal and adult male rats. Reprod Toxicol 18: 53-61. http://dx.doi.org/10.1016/j.reprotox.2003.10.011

Borch, J; Metzdorff, SB; Vinggaard, AM; Brokken, L; Dalgaard, M. (2006b). Mechanisms underlying the anti-androgenic effects of diethylhexyl phthalate in fetal rat testis. Toxicology 223: 144-155. http://dx.doi.org/10.1016/j.tox.2006.03.015

Clewell, RA; Kremer, JJ; Williams, CC; Campbell, JL; Sochaski, MA; Andersen, ME; Borghoff, SJ. (2009). Kinetics of selected di-n-butyl phthalate metabolites and fetal testosterone following repeated and single administration in pregnant rats. Toxicology 255: 80-90. http://dx.doi.org/10.1016/j.tox.2008.10.010

Culty, M; Thuillier, R; Li, W; Wang, Y; Martinez-Arguelles, D; Benjamin, C; Triantafilou, K; Zirkin, B; Papadopoulos, V. (2008). In utero exposure to di-(2-ethylhexyl) phthalate exerts both short-term and long-lasting suppressive effects on testosterone production in the rat. Biol Reprod 78: 1018-1028. http://dx.doi.org/10.1095/biolreprod.107.065649

Do, RP; Stahlhut, RW; Ponzi, D; Vom Saal, FS; Taylor, JA. (2012). Non-monotonic dose effects of in utero exposure to di(2-ethylhexyl) phthalate (DEHP) on testicular and serum testosterone and anogenital distance in male mouse fetuses. Reprod Toxicol 34: 614-621. http://dx.doi.org/10.1016/j.reprotox.2012.09.006

Drake, AJ; van den Driesche, S; Scott, HM; Hutchison, GR; Seckl, JR; Sharpe, RM. (2009). Glucocorticoids amplify dibutyl phthalate-induced disruption of testosterone production and male reproductive development. Endocrinology 150: 5055-5064. http://dx.doi.org/10.1210/en.2009-0700

Faustman, EM; Allen, BC; Kavlock, RJ; Kimmel, CA. (1994). Dose-response assessment for developmental toxicity: I characterization of data base and determination of no observed adverse effect levels. Fundam Appl Toxicol 23: 478-486. https://dx.doi.org/10.1006/faat.1994.1132

Furr, JR; Lambright, CS; Wilson, VS; Foster, PM; Gray, LE, Jr. (2014). A short-term in vivo screen using fetal testosterone production, a key event in the phthalate adverse outcome pathway, to predict disruption of sexual differentiation. Toxicol Sci 140: 403-424. https://dx.doi.org/10.1093/toxsci/kfu081

Gaido, KW; Hensley, JB; Liu, D; Wallace, DG; Borghoff, S; Johnson, KJ; Hall, SJ; Boekelheide, K. (2007). Fetal mouse phthalate exposure shows that gonocyte multinucleation is not associated with decreased testicular testosterone. Toxicol Sci 97: 491-503. http://dx.doi.org/10.1093/toxsci/kfm049

Ge, RS; Chen, GR; Dong, Q; Akingbemi, B; Sottas, CM; Santos, M; Sealfon, SC; Bernard, DJ; Hardy, MP. (2007). Biphasic effects of postnatal exposure to diethylhexylphthalate on the timing of puberty in male rats. J Androl 28: 513-520. http://dx.doi.org/10.2164/jandrol.106.001909

Giribabu, N; Sainath, SB; Reddy, PS. (2014). Prenatal di-n-butyl phthalate exposure alters reproductive functions at adulthood in male rats. Environ Toxicol 29: 534-544. http://dx.doi.org/10.1002/tox.21779

Gray, L; Barlow, N; Howdeshell, K; Ostby, J; Furr, J; Gray, C. (2009). Transgenerational effects of Di (2-ethylhexyl) phthalate in the male CRL:CD(SD) rat: Added value of assessing multiple offspring per litter. Toxicol Sci 110: 411-425. http://dx.doi.org/10.1093/toxsci/kfp109

Gray, LE; Furr, J; Tatum-Gibbs, KR; Lambright, C; Sampson, H; Hannas, BR; Wilson, VS; Hotchkiss, A; Foster, PM. (2016). Establishing the "Biological Relevance" of Dipentyl Phthalate Reductions in Fetal Rat Testosterone Production and Plasma and Testis Testosterone Levels. Toxicol Sci 149: 178-191. https://dx.doi.org/10.1093/toxsci/kfv224

Gray, LE, Jr.; Lambright, CS; Conley, JM; Evans, N; Furr, JR; Hannas, BR; Wilson, VS; Sampson, H; Foster, PMD. (2021). Genomic and hormonal biomarkers of phthalate-induced male rat reproductive developmental toxicity, Part II: A targeted RT-qPCR array approach that defines a unique adverse outcome pathway. Toxicol Sci 182: 195-214. https://dx.doi.org/10.1093/toxsci/kfab053

Guo, J; Li, XW; Liang, Y; Ge, Y; Chen, X; Lian, QQ; Ge, RS. (2013). The increased number of Leydig cells by di(2-ethylhexyl) phthalate comes from the differentiation of stem cells into Leydig cell lineage in the adult rat testis. Toxicology 306: 9-15. http://dx.doi.org/10.1016/j.tox.2013.01.021

Hannas, BR; Lambright, CS; Furr, J; Howdeshell, KL; Wilson, VS; Gray, LE. (2011). Dose-response assessment of fetal testosterone production and gene expression levels in rat testes following in utero exposure to diethylhexyl phthalate, diisobutyl phthalate, diisoheptyl phthalate, and diisononyl phthalate. Toxicol Sci 123: 206-216. http://dx.doi.org/10.1093/toxsci/kfr146

Higuchi, TT; Palmer, JS; Gray, LE, Jr; Veeramachaneni, DN. (2003). Effects of dibutyl phthalate in male rabbits following in utero, adolescent, or postpubertal exposure. Toxicol Sci 72: 301-313. http://dx.doi.org/10.1093/toxsci/kfg036

Hoshino, N; Iwai, M; Okazaki, Y. (2005). A two-generation reproductive toxicity study of dicyclohexyl phthalate in rats. J Toxicol Sci 30: 79-96. http://dx.doi.org/10.2131/jts.30.s79

Howdeshell, KL; Furr, J; Lambright, CR; Rider, CV; Wilson, VS; Gray, LE, Jr. (2007). Cumulative effects of dibutyl phthalate and diethylhexyl phthalate on male rat reproductive tract development: Altered fetal steroid hormones and genes. Toxicol Sci 99: 190-202. http://dx.doi.org/10.1093/toxsci/kfm069

Howdeshell, KL; Rider, CV; Wilson, VS; Furr, JR; Lambright, CR; Gray, LE. (2015). Dose addition models based on biologically relevant reductions in fetal testosterone accurately predict postnatal reproductive tract alterations by a phthalate mixture in rats. Toxicol Sci 148: 488-502. https://dx.doi.org/10.1093/toxsci/kfv196

Howdeshell, KL; Wilson, VS; Furr, J; Lambright, CR; Rider, CV; Blystone, CR; Hotchkiss, AK; Gray, LE, Jr. (2008). A mixture of five phthalate esters inhibits fetal testicular testosterone production in the Sprague-Dawley rat in a cumulative, dose-additive manner. Toxicol Sci 105: 153-165. https://dx.doi.org/10.1093/toxsci/kfn077

Johnson, KJ; Hensley, JB; Kelso, MD; Wallace, DG; Gaido, KW. (2007). Mapping gene expression changes in the fetal rat testis following acute dibutyl phthalate exposure defines a complex temporal cascade of responding cell types. Biol Reprod 77: 978-989. http://dx.doi.org/10.1095/biolreprod.107.062950

Johnson, KJ; McDowell, EN; Viereck, MP; Xia, JQ. (2011). Species-specific dibutyl phthalate fetal testis endocrine disruption correlates with inhibition of SREBP2-dependent gene expression pathways. Toxicol Sci 120: 460-474. https://dx.doi.org/10.1093/toxsci/kfr020

Kim, TS; Jung, KK; Kim, SS; Kang, IH; Baek, JH; Nam, HS; Hong, SK; Lee, BM; Hong, JT; Oh, KW; Kim, HS; Han, SY; Kang, TS. (2010). Effects of in utero exposure to DI(n-Butyl) phthalate on development of male reproductive tracts in Sprague-Dawley rats. J Toxicol Environ Health A 73: 1544-1559. http://dx.doi.org/10.1080/15287394.2010.511579

Klinefelter, GR; Laskey, JW; Winnik, WM; Suarez, JD; Roberts, NL; Strader, LF; Riffle, BW; Veeramachaneni, DN. (2012). Novel molecular targets associated with testicular dysgenesis induced by gestational exposure to diethylhexyl phthalate in the rat: a role for estradiol. Reproduction 144: 747-761. http://dx.doi.org/10.1530/REP-12-0266

Kuhl, AJ; Ross, SM; Gaido, KW. (2007). CCAAT/enhancer binding protein beta, but not steroidogenic factor-1, modulates the phthalate-induced dysregulation of rat fetal testicular steroidogenesis. Endocrinology 148: 5851-5864. http://dx.doi.org/10.1210/en.2007-0930

Kurahashi, N; Kondo, T; Omura, M; Umemura, T; Ma, M; Kishi, R. (2005). The effects of subacute inhalation of di (2-ethylhexyl) phthalate (DEHP) on the testes of prepubertal Wistar rats. J Occup Health 47: 437-444. http://dx.doi.org/10.1539/joh.47.437

Lehmann, KP; Phillips, S; Sar, M; Foster, PMD; Gaido, KW. (2004). Dose-dependent alterations in gene expression and testosterone synthesis in the fetal testes of male rats exposed to di (n-butyl) phthalate. Toxicol Sci 81: 60-68. http://dx.doi.org/10.1093/toxsci/kfh169

Li, H; Jiang, Y; Liu, M; Yu, J; Feng, X; Xu, X; Wang, H; Zhang, J; Sun, X; Yu, Y. (2023). DNA methylation-mediated inhibition of MGARP is involved in impaired progeny testosterone synthesis in mice exposed to DBP in utero. Environ Toxicol 38: 914-925. http://dx.doi.org/10.1002/tox.23734

Li, N; Chen, X; Zhou, X; Zhang, W; Yuan, J; Feng, J. (2015). The mechanism underlying dibutyl phthalate induced shortened anogenital distance and hypospadias in rats. J Pediatr Surg 50: 2078-2083. http://dx.doi.org/10.1016/j.jpedsurg.2015.08.046

Li, X; Chen, X; Hu, G; Li, L; Su, H; Wang, Y; Chen, D; Zhu, Q; Li, C; Li, J; Wang, M; Lian, Q; Ge, R. (2016). Effects of in utero exposure to dicyclohexyl phthalate on rat fetal leydig cells. Int J Environ Res Public Health 13: 1. http://dx.doi.org/10.3390/ijerph13030246

Li, XW; Liang, Y; Su, Y; Deng, H; Li, XH; Guo, J; Lian, QQ; Ge, RS. (2012). Adverse effects of di-(2-ethylhexyl) phthalate on Leydig cell regeneration in the adult rat testis. Toxicol Lett 215: 84-91. http://dx.doi.org/10.1016/j.toxlet.2012.10.001

Lin, H; Ge, R; Chen, G; Hu, G; Dong, L; Lian, Q; Hardy, D; Sottas, C; Li, X; Hardy, M. (2008). Involvement of testicular growth factors in fetal Leydig cell aggregation after exposure to phthalate in utero. Proc Natl Acad Sci USA 105: 7218-7222. http://dx.doi.org/10.1073/pnas.0709260105

Lin, H; Lian, Q; Hu, G; Jin, Y; Zhang, Y; Hardy, D; Chen, G; Lu, Z; Sottas, C; Hardy, M; Ge, R. (2009). In utero and lactational exposures to diethylhexyl-phthalate affect two populations of Leydig cells in male Long-Evans rats. Biol Reprod 80: 882-888. http://dx.doi.org/10.1095/biolreprod.108.072975

Lv, Y; Fang, Y; Chen, P; Duan, Y; Huang, T; Ma, L; Xie, L; Chen, X; Chen, X; Gao, J; Ge, RS. (2019). Dicyclohexyl phthalate blocks Leydig cell regeneration in adult rat testis. Toxicology 411: 60-70. https://dx.doi.org/10.1016/j.tox.2018.10.020

Ma, M; Kondo, T; Ban, S; Umemura, T; Kurahashi, N; Takeda, M; Kishi, R. (2006). Exposure of prepubertal female rats to inhaled di(2-ethylhexyl)phthalate affects the onset of puberty and postpubertal reproductive functions. Toxicol Sci 93: 164-171. http://dx.doi.org/10.1093/toxsci/kfl036

MacLeod, DJ; Sharpe, RM; Welsh, M; Fisken, M; Scott, HM; Hutchison, GR; Drake, AJ; van Den Driesche, S. (2010). Androgen action in the masculinization programming window and development of male reproductive organs. Int J Androl 33: 279-287. https://dx.doi.org/10.1111/j.1365-2605.2009.01005.x

Mahood, IK; Scott, HM; Brown, R; Hallmark, N; Walker, M; Sharpe, RM. (2007). In utero exposure to di(n-butyl) phthalate and testicular dysgenesis: Comparison of fetal and adult end points and their dose sensitivity. Environ Health Perspect 115: 55-61. http://dx.doi.org/10.1289/ehp.9366

Martino-Andrade, AJ; Morais, RN; Botelho, GG; Muller, G; Grande, SW; Carpentieri, GB; Leao, GM; Dalsenter, PR. (2008). Coadministration of active phthalates results in disruption of foetal testicular function in rats. Int J Androl 32: 704-712. http://dx.doi.org/10.1111/j.1365-2605.2008.00939.x

McKinnell, C; Mitchell, RT; Walker, M; Morris, K; Kelnar, CJH; Wallace, WH; Sharpe, RM. (2009). Effect of fetal or neonatal exposure to monobutyl phthalate (MBP) on testicular development and function in the marmoset. Hum Reprod 24: 2244-2254. http://dx.doi.org/10.1093/humrep/dep200

Moody, S; Goh, H; Bielanowicz, A; Rippon, P; Loveland, KL; Itman, C. (2013). Prepubertal mouse testis growth and maturation and androgen production are acutely sensitive to di-n-butyl phthalate. Endocrinology 154: 3460-3475. http://dx.doi.org/10.1210/en.2012-2227

Mylchreest, E; Sar, M; Wallace, DG; Foster, PMD. (2002). Fetal testosterone insufficiency and abnormal proliferation of Leydig cells and gonocytes in rats exposed to di(n-butyl) phthalate. 16: 19-28.

Nagao, T; Ohta, R; Marumo, H; Shindo, T; Yoshimura, S; Ono, H. (2000). Effect of butyl benzyl phthalate in Sprague-Dawley rats after gavage administration: A two-generation reproductive study. Reprod Toxicol 14: 513-532. https://dx.doi.org/10.1016/S0890-6238(00)00105-2

NASEM. (2017). Application of systematic review methods in an overall strategy for evaluating low-dose toxicity from endocrine active chemicals. In Consensus Study Report. Washington, D.C.: The National Academies Press. https://dx.doi.org/10.17226/24758

Pan, Y; Wang, X; Yeung, LWY; Sheng, N; Cui, Q; Cui, R; Zhang, H; Dai, J. (2017). Dietary exposure to di-isobutyl phthalate increases urinary 5-methyl-2'-deoxycytidine level and affects

reproductive function in adult male mice. J Environ Sci 61: 14-23. http://dx.doi.org/10.1016/j.jes.2017.04.036
Rajagopal, G; Bhaskaran, RS; Karundevi, B. (2019). Maternal di-(2-ethylhexyl) phthalate exposure alters hepatic insulin signal transduction and glucoregulatory events in rat F1 male offspring. J Appl Toxicol 39: 751-763. http://dx.doi.org/10.1002/jat.3764
Saillenfait, AM; Sabaté, JP; Denis, F; Antoine, G; Robert, A; Roudot, AC; Ndiaye, D; Eljarrat, E. (2017). Evaluation of the effects of α-cypermethrin on fetal rat testicular steroidogenesis. Reprod Toxicol 72: 106-114. http://dx.doi.org/10.1016/j.reprotox.2017.06.133
Saillenfait, AM; Sabaté, JP; Robert, A; Rouiller-Fabre, V; Roudot, AC; Moison, D; Denis, F. (2013). Dose-dependent alterations in gene expression and testosterone production in fetal rat testis after exposure to di-n-hexyl phthalate. J Appl Toxicol 33: 1027-1035. http://dx.doi.org/10.1002/jat.2896
Scarano, WR; Toledo, FC; Guerra, MT; Pinheiro, PFF; Domeniconi, RF; Felisbino, SL; Campos, SG; Taboga, SR; Kempinas, WG. (2010). Functional and morphological reproductive aspects in male rats exposed to di-n-butyl phthalate (DBP) in utero and during lactation. J Toxicol Environ Health A 73: 972-984. http://dx.doi.org/10.1080/15287391003751760
Schmitt, EE; Vellers, HL; Porter, WW; Lightfoot, JT. (2016). Environmental endocrine disruptor affects voluntary physical activity in mice. Med Sci Sports Exerc 48: 1251-1258. https://dx.doi.org/10.1249/MSS.0000000000000908
Spade, DJ; Bai, CY; Lambright, C; Conley, JM; Boekelheide, K; Gray, LE. (2018). Validation of an automated counting procedure for phthalate-induced testicular multinucleated germ cells. Toxicol Lett 290: 55-61. https://dx.doi.org/10.1016/j.toxlet.2018.03.018
Struve, MF; Gaido, KW; Hensley, JB; Lehmann, KP; Ross, SM; Sochaski, MA; Willson, GA; Dorman, DC. (2009). Reproductive toxicity and pharmacokinetics of di-n-butyl phthalate (DBP) following dietary exposure of pregnant rats. Birth Defects Res B Dev Reprod Toxicol 86: 345-354. http://dx.doi.org/10.1002/bdrb.20199
U.S. EPA. (2012). Benchmark dose technical guidance [EPA Report]. (EPA100R12001). Washington, DC: U.S. Environmental Protection Agency, Risk Assessment Forum. https://www.epa.gov/risk/benchmark-dose-technical-guidance
U.S. EPA. (2023). Draft Proposed Approach for Cumulative Risk Assessment of High-Priority Phthalates and a Manufacturer-Requested Phthalate under the Toxic Substances Control Act. (EPA-740-P-23-002). Washington, DC: U.S. Environmental Protection Agency, Office of Chemical Safety and Pollution Prevention. https://www.regulations.gov/document/EPA-HQ-OPPT-2022-0918-0009
U.S. EPA. (2025a). Non-Cancer Human Health Hazard Assessment for Butyl Benzyl Phthalate (BBP). Washington, DC: Office of Pollution Prevention and Toxics.
U.S. EPA. (2025b). Non-Cancer Human Health Hazard Assessment for Dibutyl Phthalate (DBP). Washington, DC: Office of Pollution Prevention and Toxics.
U.S. EPA. (2025c). Non-Cancer Human Health Hazard Assessment for Dicyclohexyl Phthalate (DCHP). Washington, DC: Office of Pollution Prevention and Toxics.
U.S. EPA. (2025d). Non-Cancer Human Health Hazard Assessment for Diisobutyl Phthalate (DIBP). Washington, DC: Office of Pollution Prevention and Toxics.
U.S. EPA. (2025e). Non-Cancer Human Health Hazard Assessment for Diisononyl Phthalate (DINP). (EPA-740-R-25-009). Washington, DC: Office of Pollution Prevention and Toxics. https://www.regulations.gov/document/EPA-HQ-OPPT-2018-0436-0137
U.S. EPA. (2025f). Risk Evaluation for Butyl Benzyl Phthalate (BBP). Washington, DC: Office of Pollution Prevention and Toxics.
U.S. EPA. (2025g). Risk Evaluation for Dibutyl Phthalate (DBP). Washington, DC: Office of Pollution Prevention and Toxics.

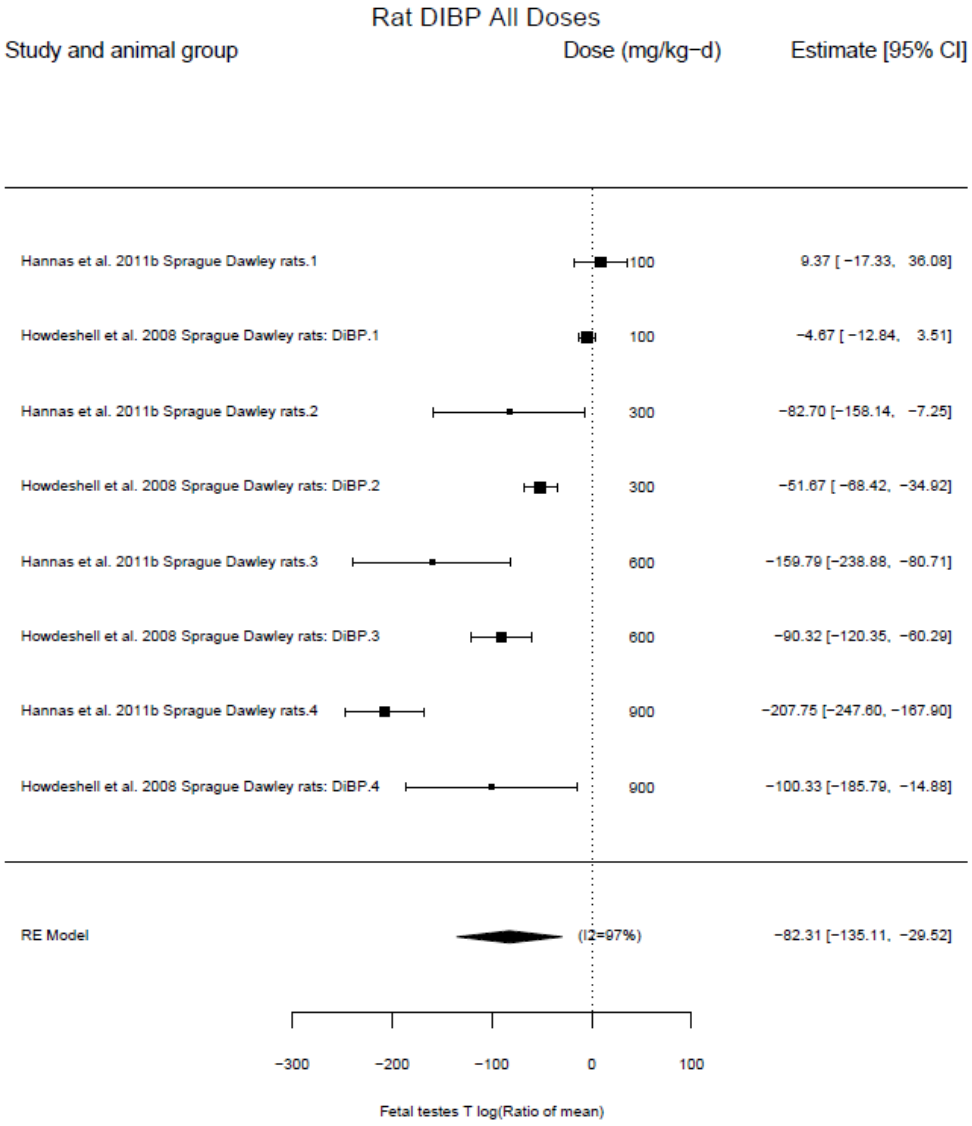
U.S. EPA. (2025h). Risk Evaluation for Dicyclohexyl Phthalate (DCHP). Washington, DC: Office of Pollution Prevention and Toxics.
U.S. EPA. (2025i). Risk Evaluation for Diethylhexyl Phthalate (DEHP). Washington, DC: Office of Pollution Prevention and Toxics.
U.S. EPA. (2025j). Risk Evaluation for Diisobutyl Phthalate (DIBP). Washington, DC: Office of Pollution Prevention and Toxics.
U.S. EPA. (2025k). Science Advisory Committee on Chemicals (SACC) meeting minutes and final report - Peer Review of the Draft Risk Evaluations of Dibutyl phthalate (DBP), Di(2-ethylhexyl) phthalate (DEHP), and Dicyclohexyl phthalate (DCHP), and the Technical Support Documents for Butylbenzyl phthalate (BBP) and Diisobutyl phthalate (DIBP). Washington, DC. https://www.regulations.gov/docket/EPA-HQ-OPPT-2024-0551
U.S. EPA. (2025l). Systematic Review Protocol for Dicyclohexyl Phthalate (DCHP). Washington, DC: Office of Pollution Prevention and Toxics.
U.S. EPA. (2025m). Technical Support Document for the Cumulative Risk Analysis of Di(2-ethylhexyl) Phthalate (DEHP), Dibutyl Phthalate (DBP), Butyl Benzyl Phthalate (BBP), Diisobutyl Phthalate (DIBP), Dicyclohexyl Phthalate (DCHP), and Diisononyl Phthalate (DINP) Under the Toxic Substances Control Act (TSCA). Washington, DC: Office of Pollution Prevention and Toxics.
van den Driesche, S; Kolovos, P; Platts, S; Drake, AJ; Sharpe, RM. (2012). Inter-relationship between testicular dysgenesis and Leydig cell function in the masculinization programming window in the rat. PLoS ONE 7: e30111. http://dx.doi.org/10.1371/journal.pone.0030111
Vo, T; Jung, E; Dang, V; Jung, K; Baek, J; Choi, K; Jeung, E. (2009a). Differential effects of flutamide and di-(2-ethylhexyl) phthalate on male reproductive organs in a rat model. J Reprod Dev 55: 400-411. http://dx.doi.org/10.1262/jrd.20220
Vo, TTB; Jung, EM; Dang, VH; Yoo, YM; Choi, KC; Yu, FH; Jeung, EB. (2009b). Di-(2 ethylhexyl) phthalate and flutamide alter gene expression in the testis of immature male rats. 7: 104. https://link.springer.com/article/10.1186/1477-7827-7-104
Wang, X; Sheng, N; Cui, R; Zhang, H; Wang, J; Dai, J. (2017). Gestational and lactational exposure to di-isobutyl phthalate via diet in maternal mice decreases testosterone levels in male offspring. Chemosphere 172: 260-267. http://dx.doi.org/10.1016/j.chemosphere.2017.01.011
Wilson, VS; Lambright, C; Furr, J; Ostby, J; Wood, C; Held, G; Gray, LE, Jr. (2004). Phthalate ester-induced gubernacular lesions are associated with reduced insl3 gene expression in the fetal rat testis. Toxicol Lett 146: 207-215. https://dx.doi.org/10.1016/j.toxlet.2003.09.012
Xiao-Feng, Z; Nai-Qiang, Q; Jing, Z; Zi, L; Yang, Z. (2009). Di (n-butyl) phthalate inhibits testosterone synthesis through a glucocorticoid-mediated pathway in rats. Int J Toxicol 28: 448-456. http://dx.doi.org/10.1177/1091581809342596

APPENDICES

Appendix A SUPPORTING MATERIALS FOR THE META-ANALYSIS AND BMD ANALYSIS OF FETAL TESTICULAR TESTOSTERONE IN RATS

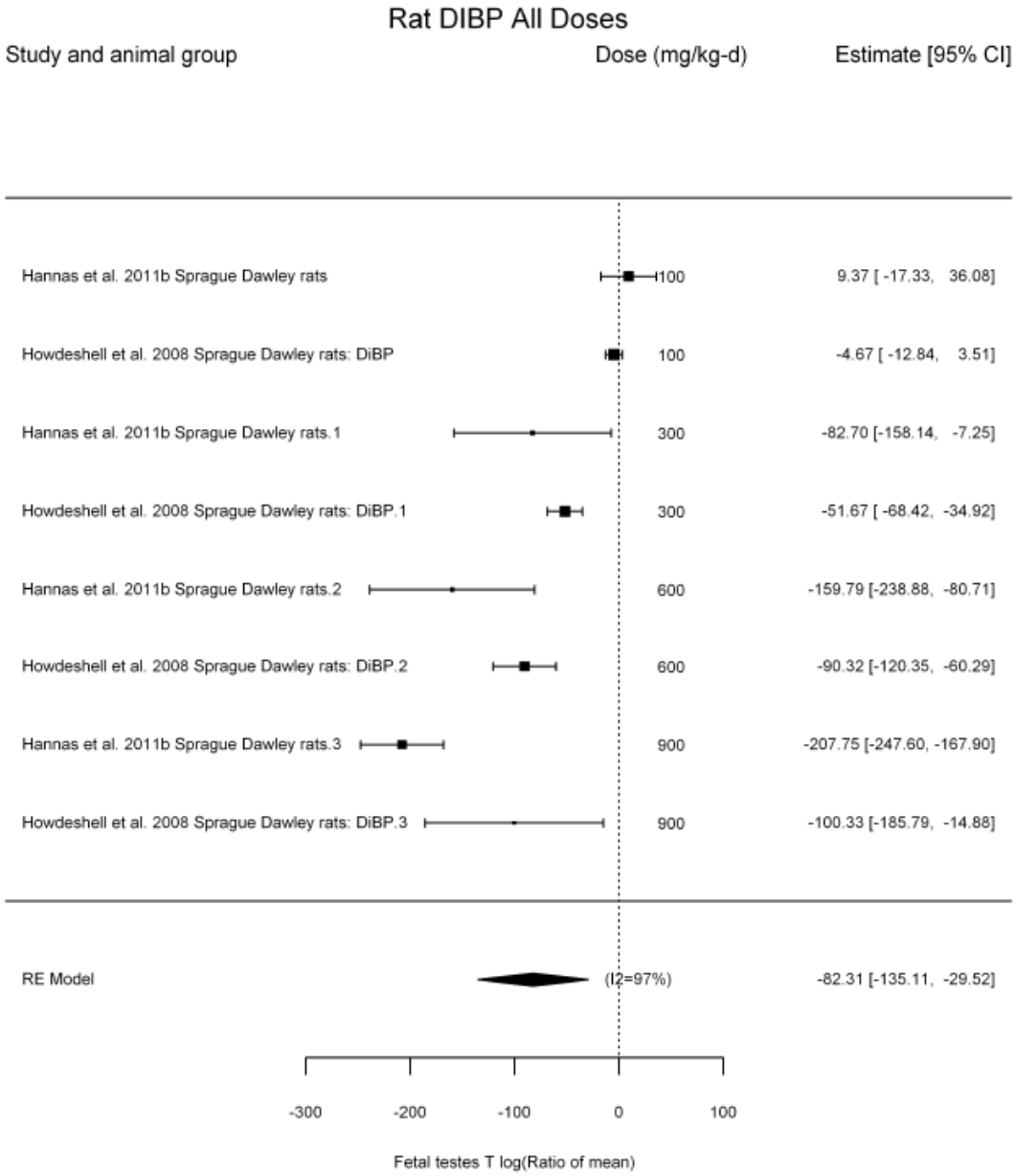
The measured outcome of free testes T log transformed ratio of means was converted to a percent change, as described in Section C-6 of NASEM (2017). In the plots below in Appendices A.1 through A.6, 5, 10 and 40 percent changes are shown as the equivalent log transformed ratio of means (*i.e.*, BMRs of −5.1, −11 and −51, respectively).

A.1 Replication of NASEM 2017 Results for Fetal Testosterone in Rats for DIBP



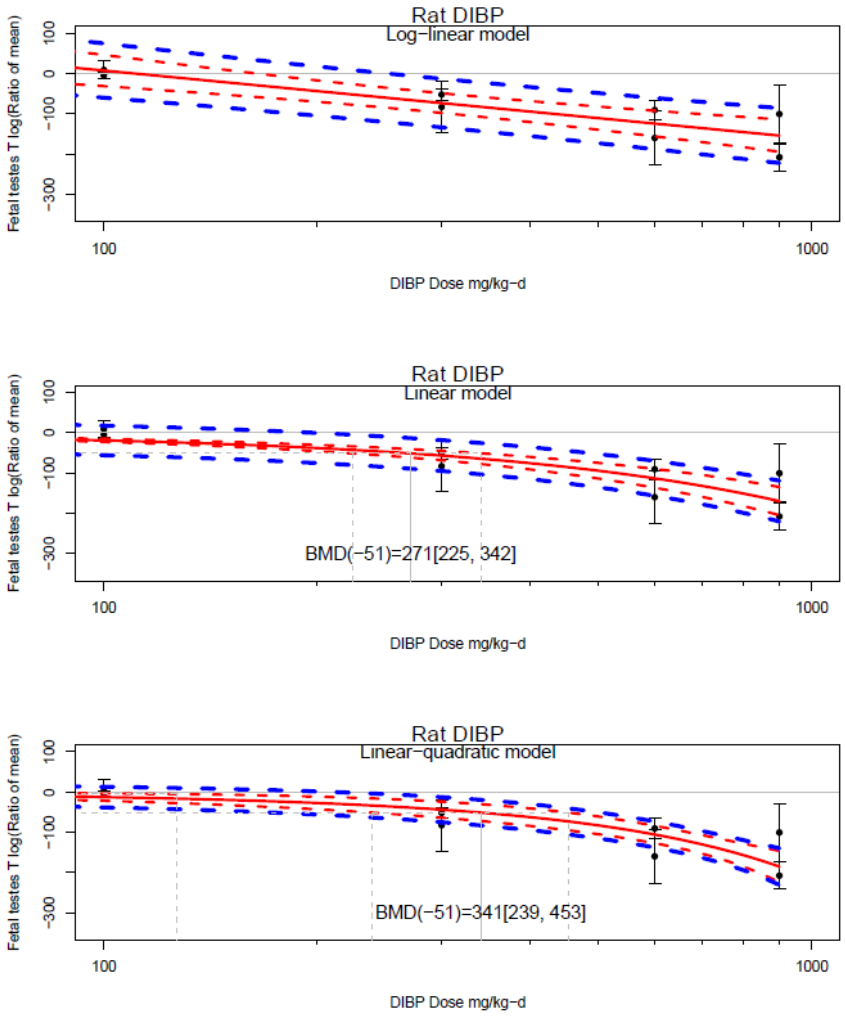
Figure_Apx A-1. Replication of NASEM (2017) Meta-analysis of Studies of DIBP and Fetal Testosterone in Rats Using Metafor Version 2.0.0

'Estimate [95% CI]' indicates the estimated effect of DIBP on free testes testosterone expressed as the log transformed ratio of means.

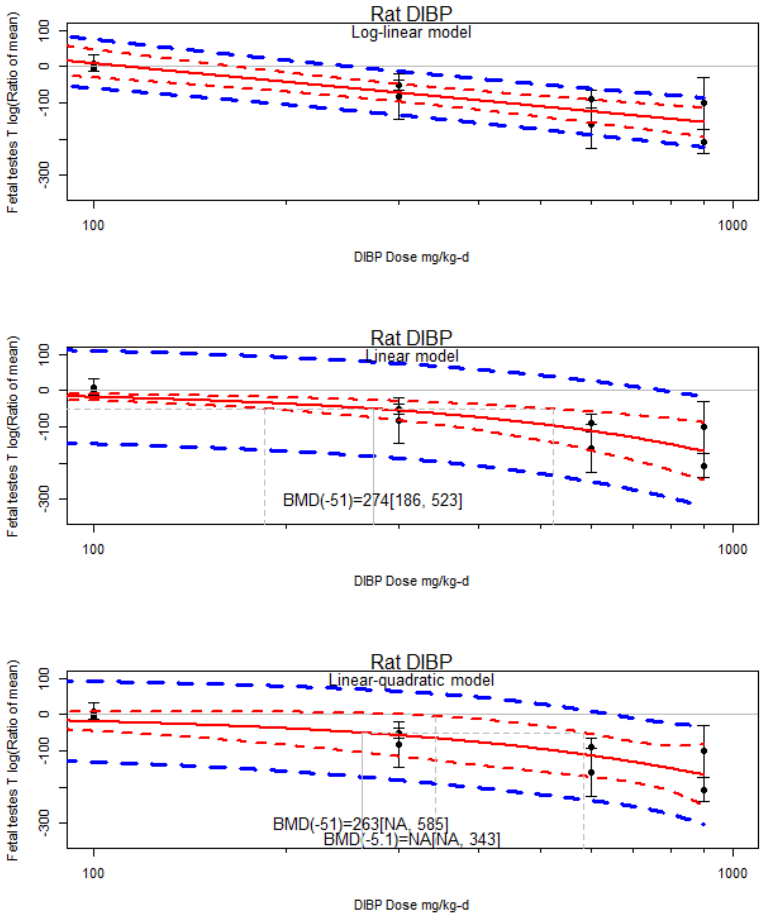


Figure_Apx A-2. Replication of NASEM (2017) Meta-analysis of Studies of DIBP and Fetal Testosterone in Rats Using Metafor Version 4.6.0

'Estimate [95% CI]' indicates the estimated effect of DIBP on free testes testosterone expressed as the log transformed ratio of means.

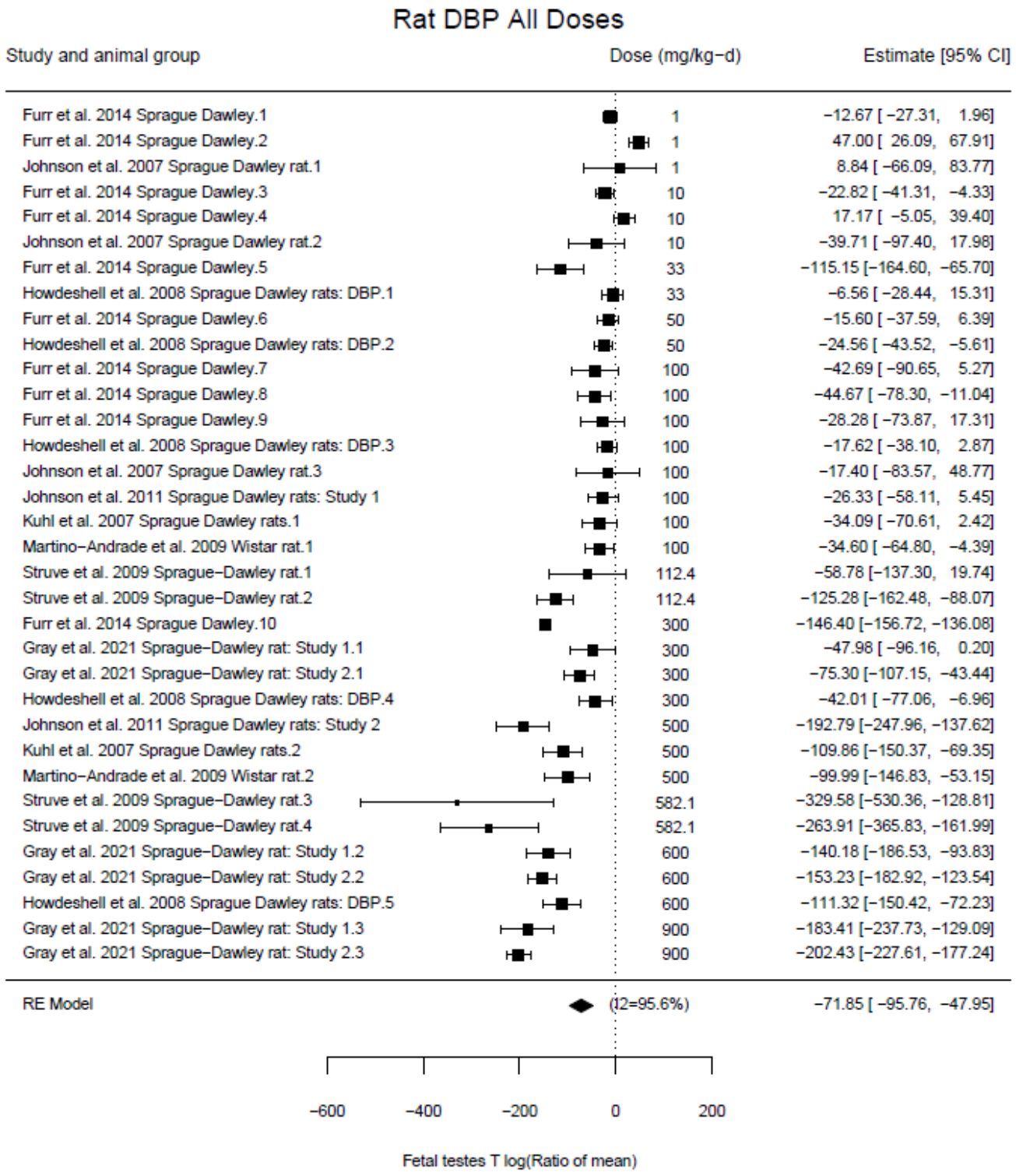


Figure_Apx A-3. Replication of NASEM (2017) Results: Benchmark Dose Estimates from Rat Studies of DIBP and Fetal Testosterone (Metafor Version 2.0.0)



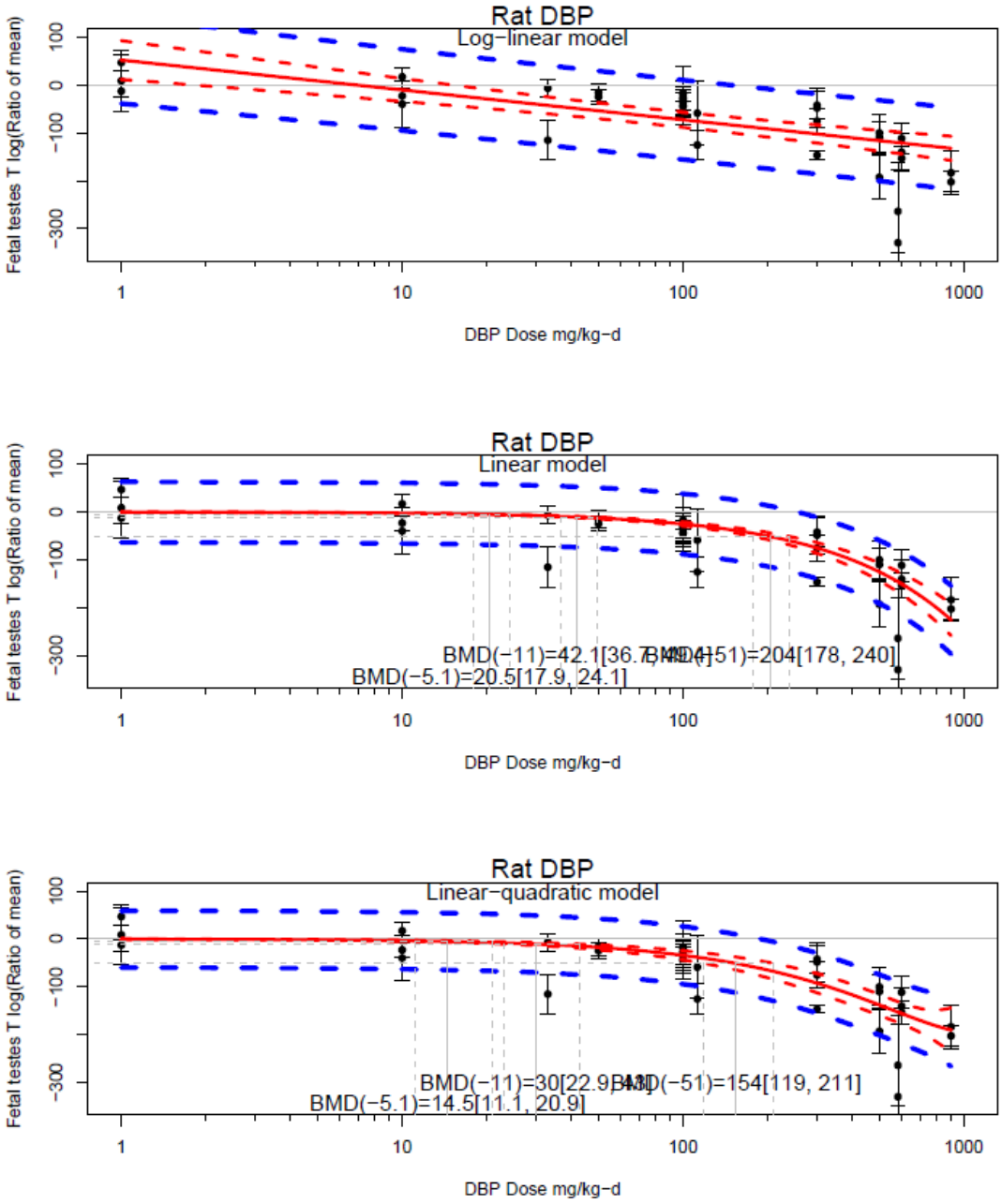
Figure_Apx A-4. Replication of NASEM (2017) Results: Benchmark Dose Estimates from Rat Studies of DIBP and Fetal Testosterone (Metafor Version 4.6.0)

A.2 Dibutyl Phthalate (DBP) – Updated Analysis

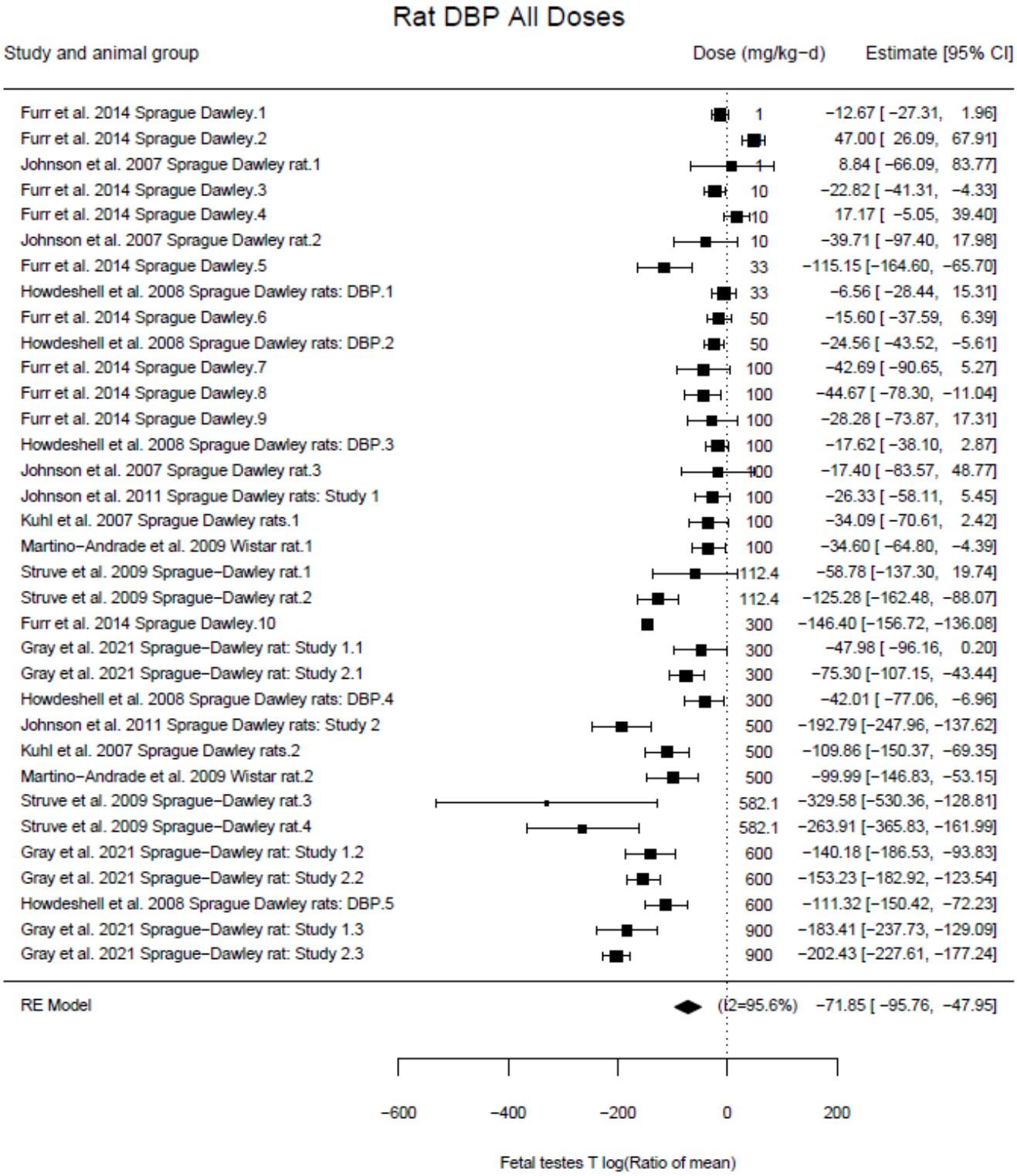


Figure_Apx A-5. Updated Meta-analysis of Studies of DBP and Fetal Testosterone in Rats (Metafor Version 2.0.0)

'Estimate [95% CI]' indicates the estimated effect of DBP on free testes testosterone expressed as the log transformed ratio of means.

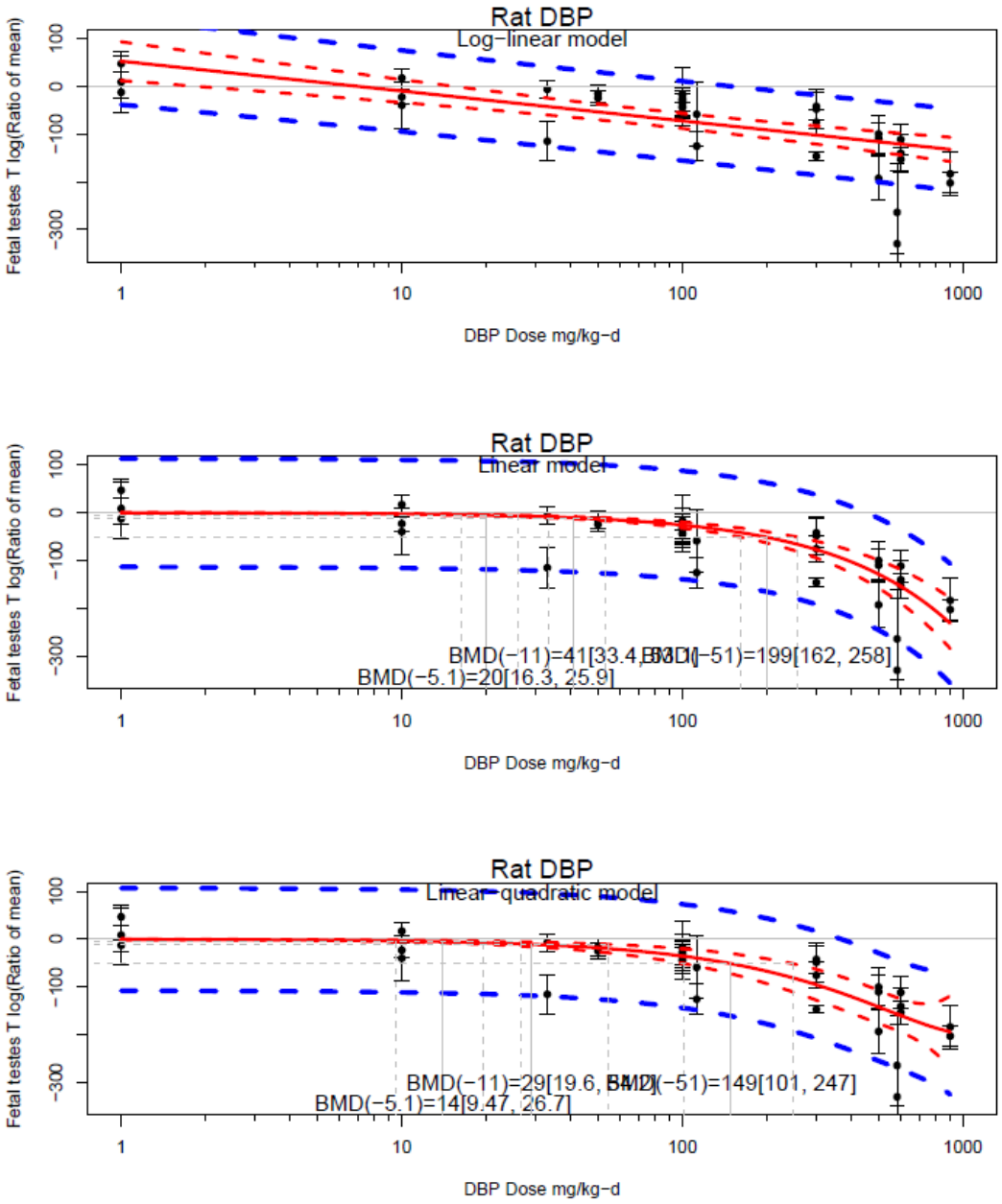


Figure_Apx A-6. Updated Benchmark Dose Estimates from Rat Studies of DBP and Fetal Testosterone (Metafor Version 2.0.0)



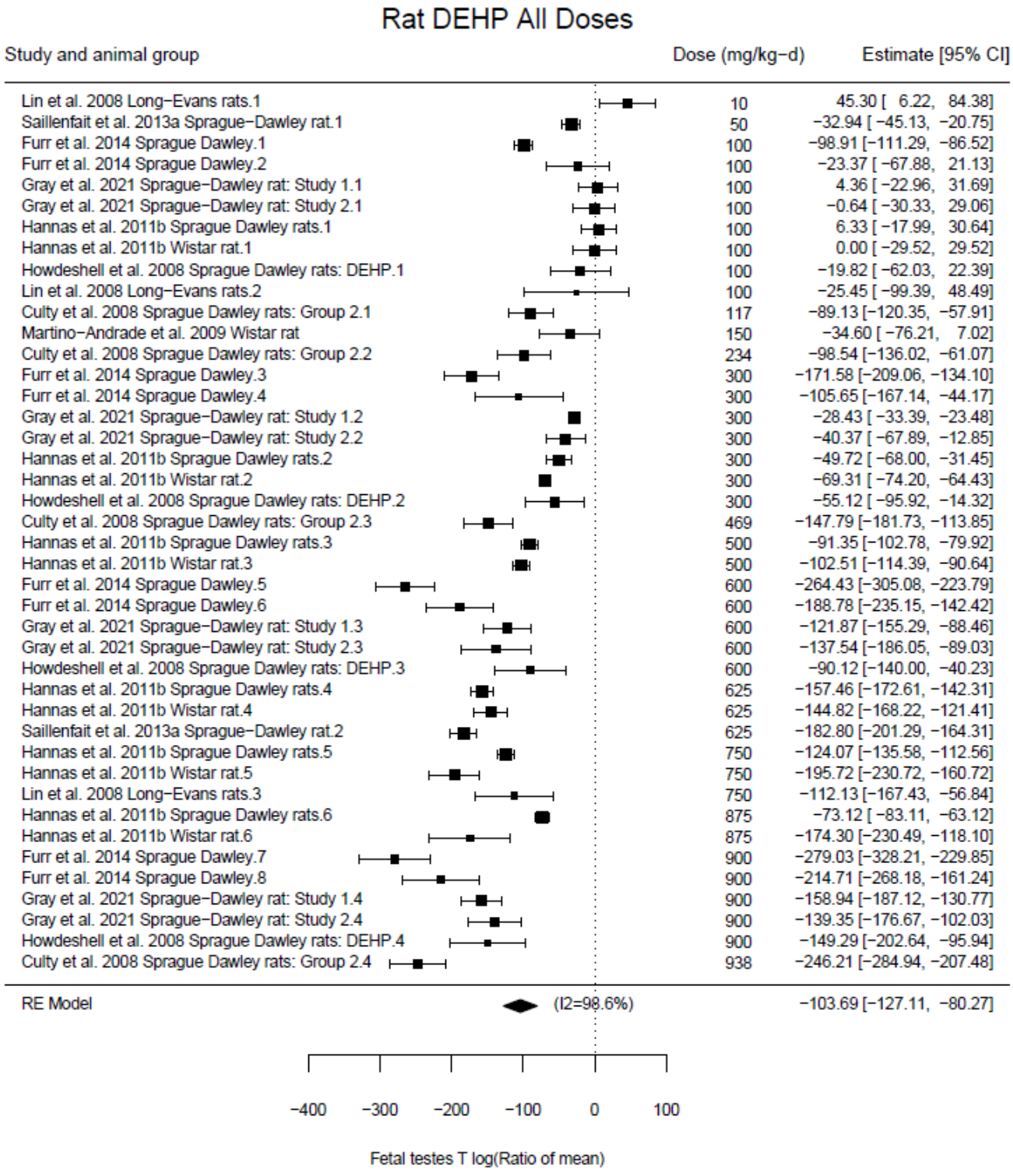
Figure_Apx A-7. Updated Meta-analysis of Studies of DBP and Fetal Testosterone in Rats (Metafor Version 4.6.0)

'Estimate [95% CI]' indicates the estimated effect of DBP on free testes testosterone expressed as the log transformed ratio of means.



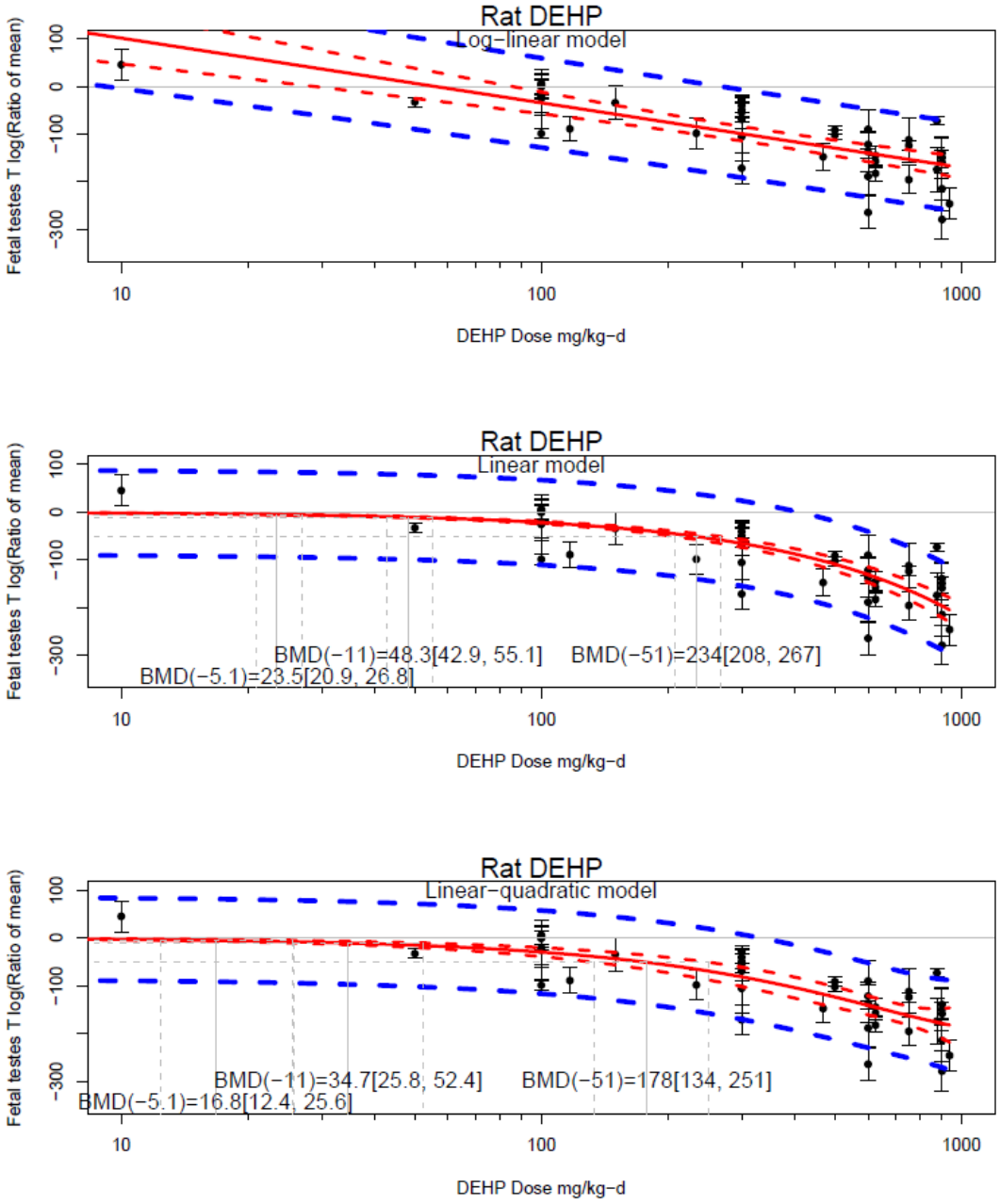
Figure_Apx A-8. Updated Benchmark Dose Estimates from Rat Studies of DBP and Fetal Testosterone (Metafor Version 4.6.0)

A.3 Di(2-ethylhexyl) Phthalate (DEHP) – Updated Analysis

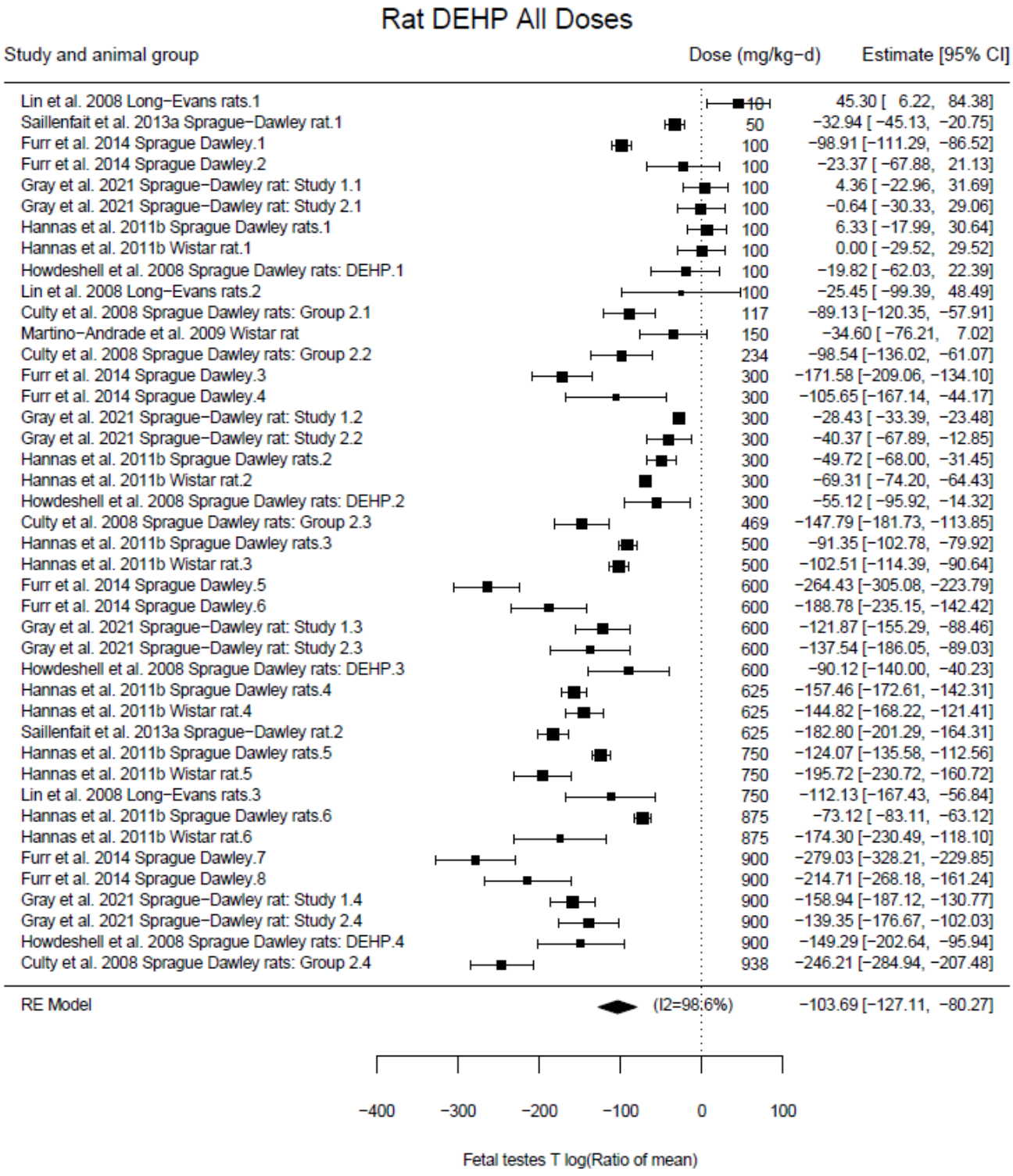


Figure_Apx A-9. Updated Meta-analysis of Studies of DEHP and Fetal Testosterone in Rats (Metafor Version 2.0.0)

'Estimate [95% CI]' indicates the estimated effect of DEHP on free testes testosterone expressed as the log transformed ratio of means.

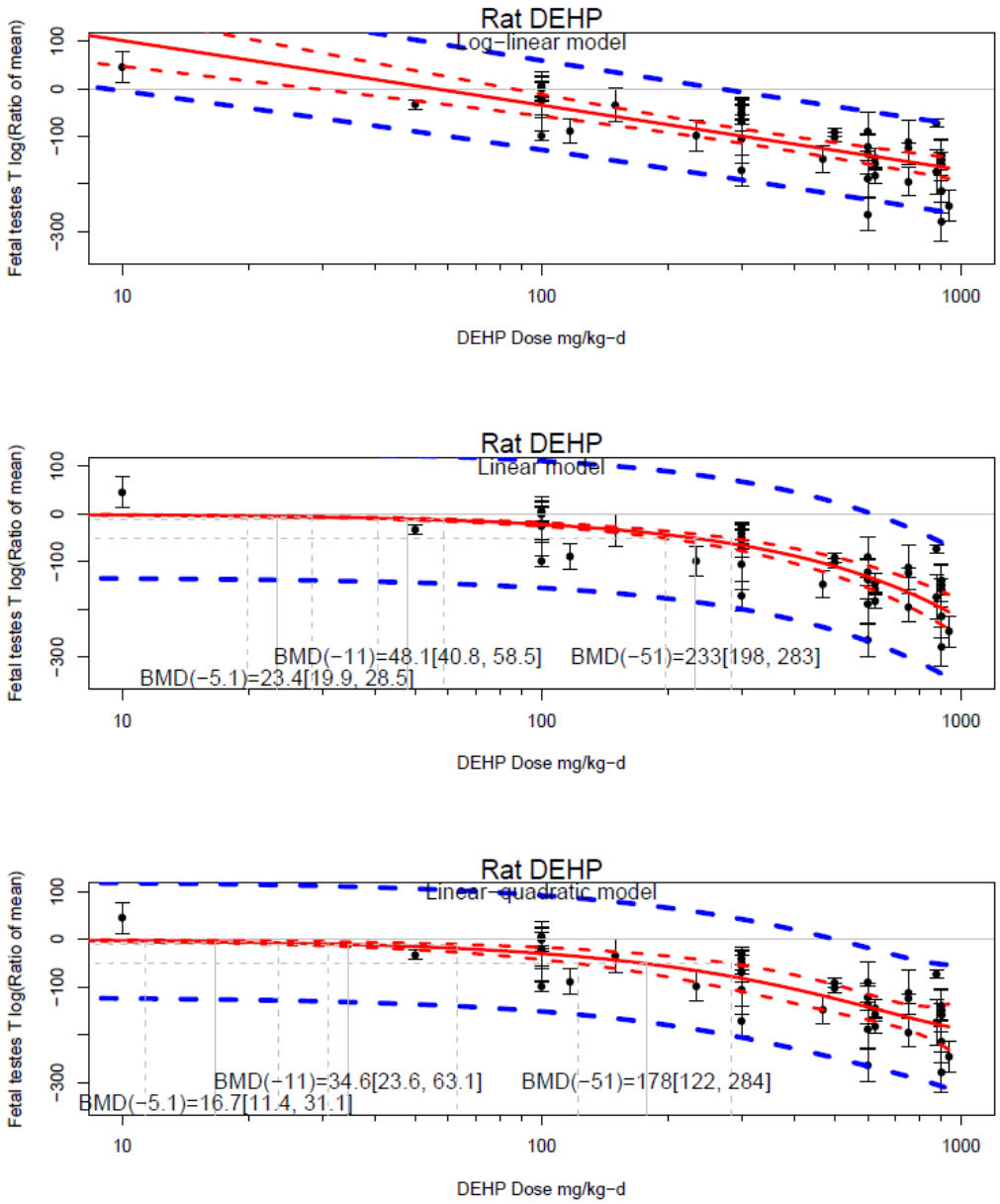


Figure_Apx A-10. Updated Benchmark Dose Estimates from Rat Studies of DEHP and Fetal Testosterone (Metafor Version 2.0.0)



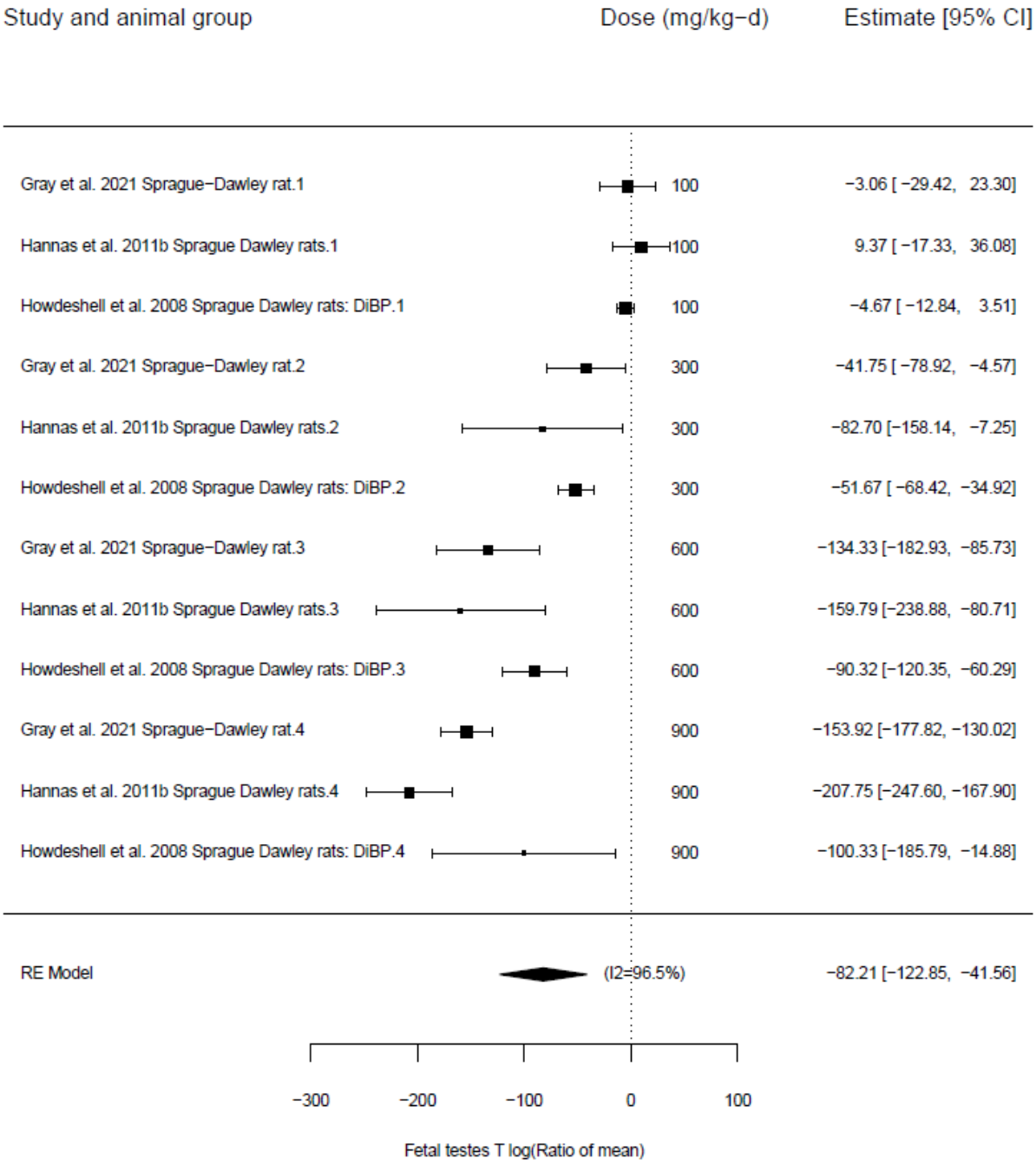
Figure_Apx A-11. Updated Meta-analysis of Studies of DEHP and Fetal Testosterone in Rats (Metafor Version 4.6.0)

'Estimate [95% CI]' indicates the estimated effect of DEHP on free testes testosterone expressed as the log transformed ratio of means.



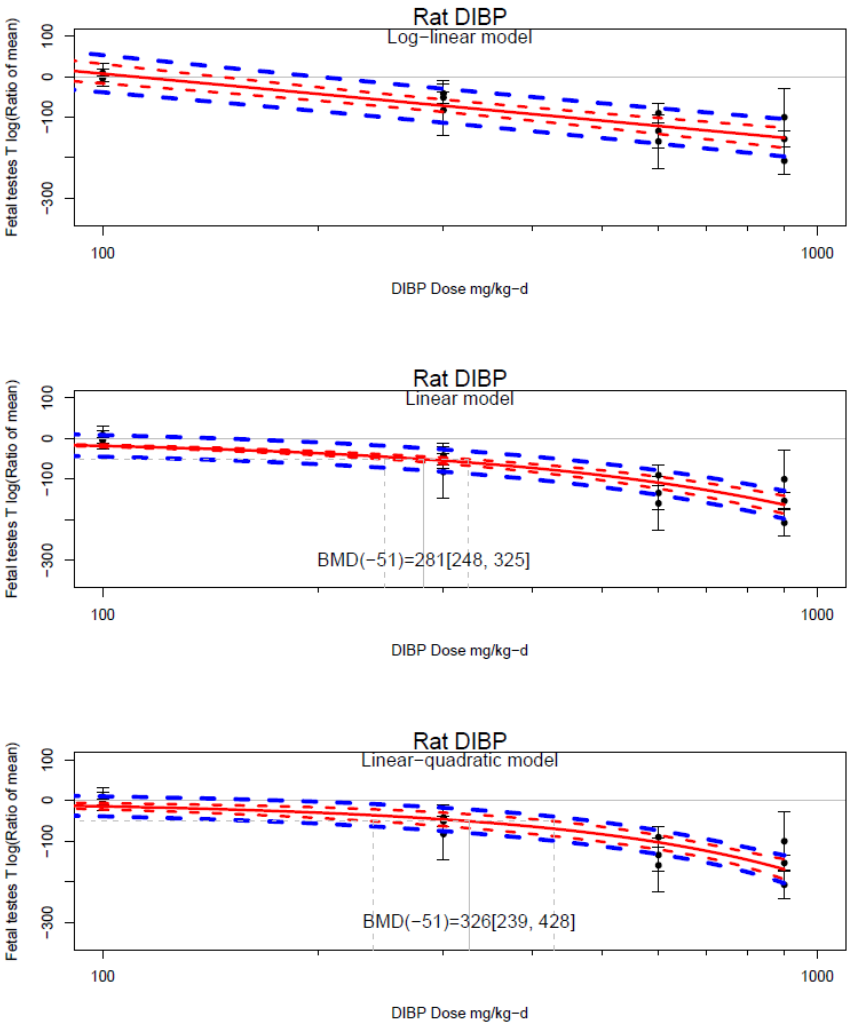
Figure_Apx A-12. Updated Benchmark Dose Estimates from Rat Studies of DEHP and Fetal Testosterone (Metafor Version 4.6.0)

A.4 Diisobutyl Phthalate (DIBP) – Updated Analysis

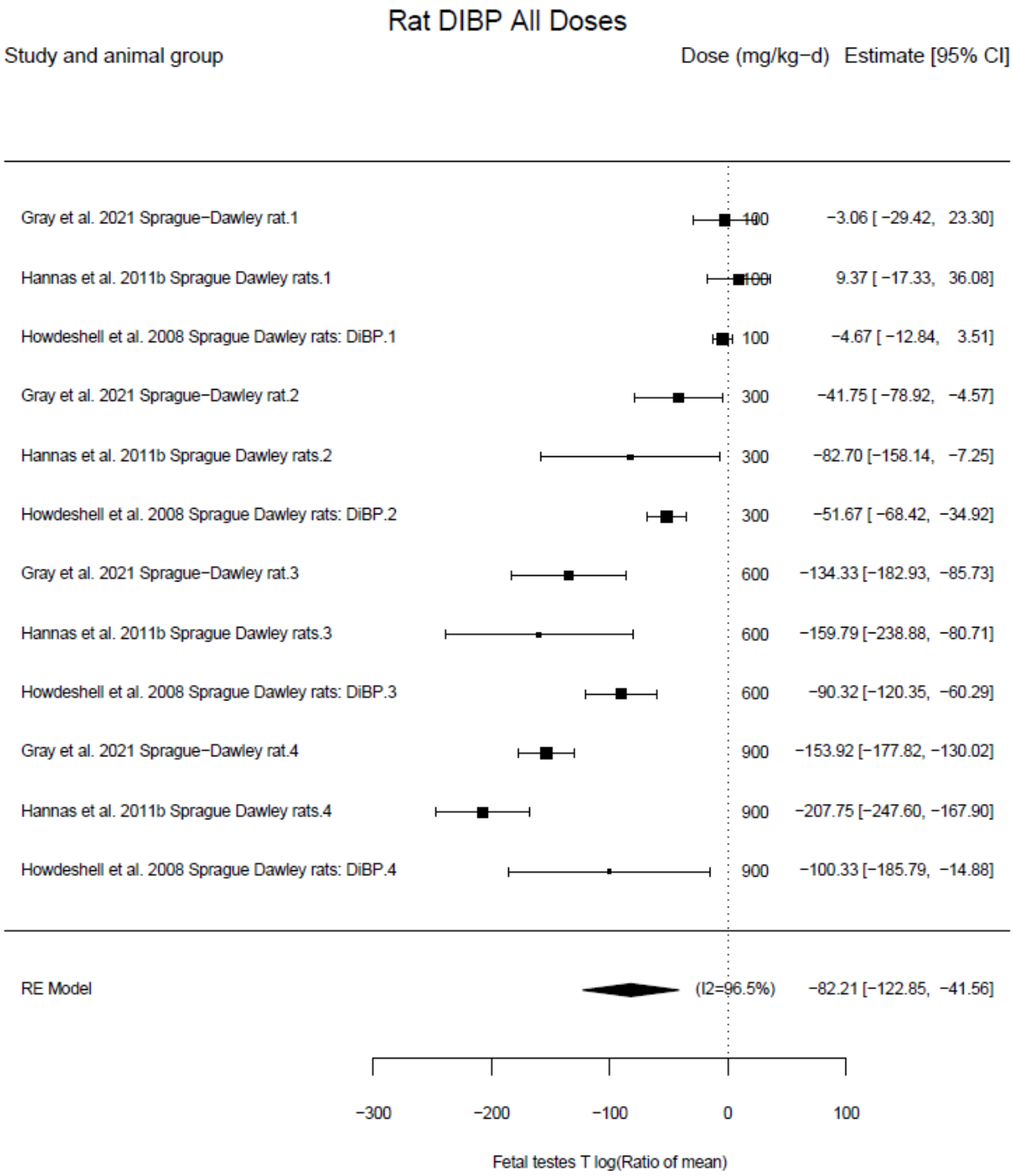


Figure_Apx A-13. Updated Meta-analysis of Studies of DIBP and Fetal Testosterone in Rats (Metafor Version 2.0.0)

'Estimate [95% CI]' indicates the estimated effect of DIBP on free testes testosterone expressed as the log transformed ratio of means.

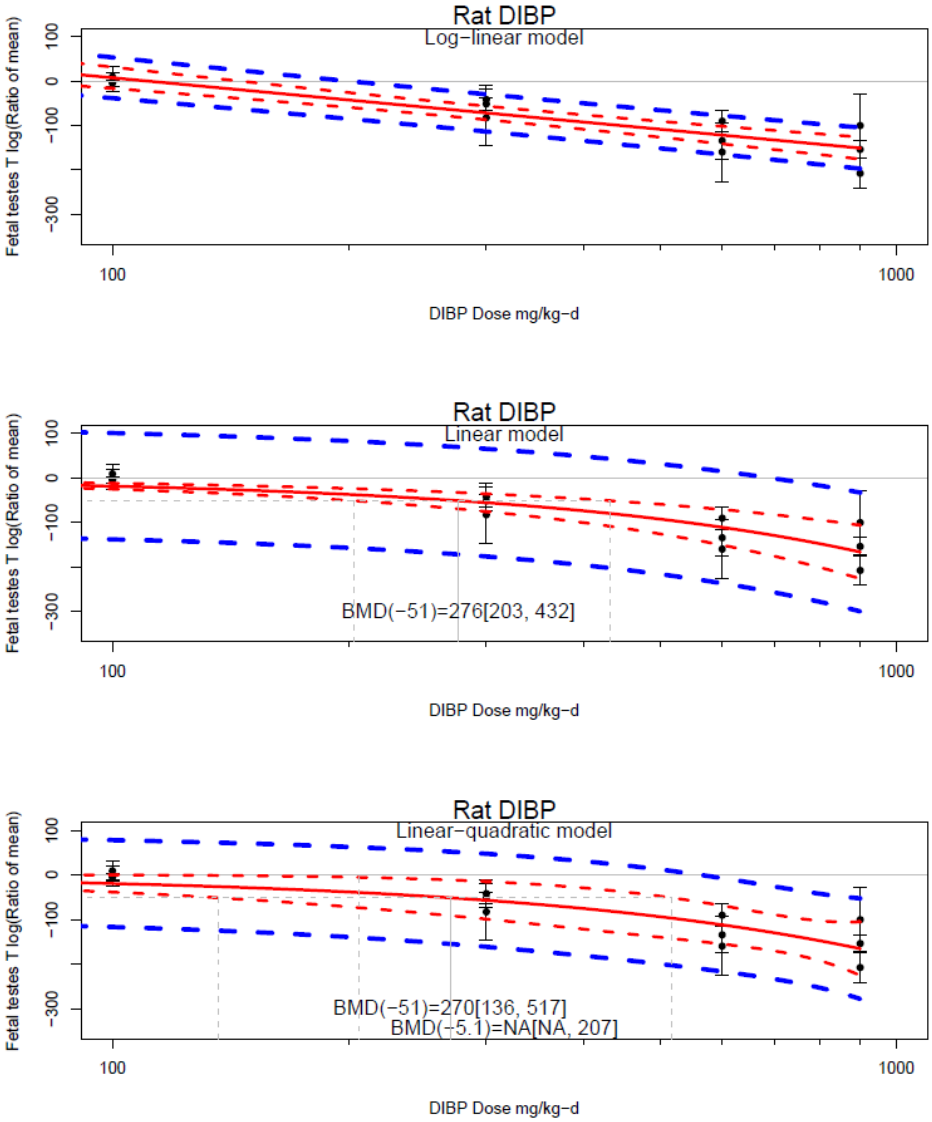


Figure_Apx A-14. Updated Benchmark Dose Estimates from Rat Studies of DIBP and Fetal Testosterone (Metafor Version 2.0.0)



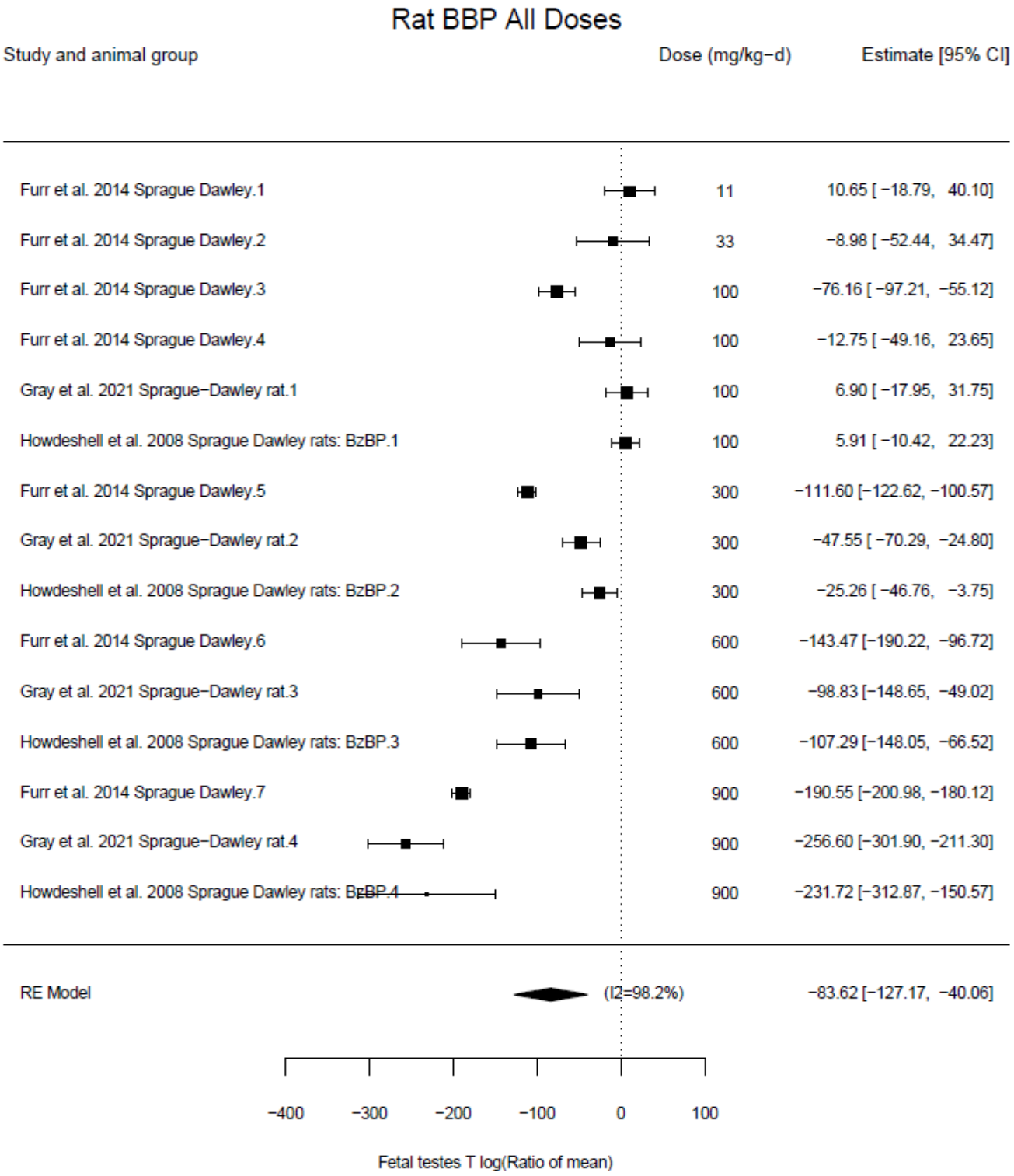
Figure_Apx A-15. Updated Meta-analysis of Studies of DIBP and Fetal Testosterone in Rats (Metafor Version 4.6.0)

'Estimate [95% CI]' indicates the estimated effect of DIBP on free testes testosterone expressed as the log transformed ratio of means.



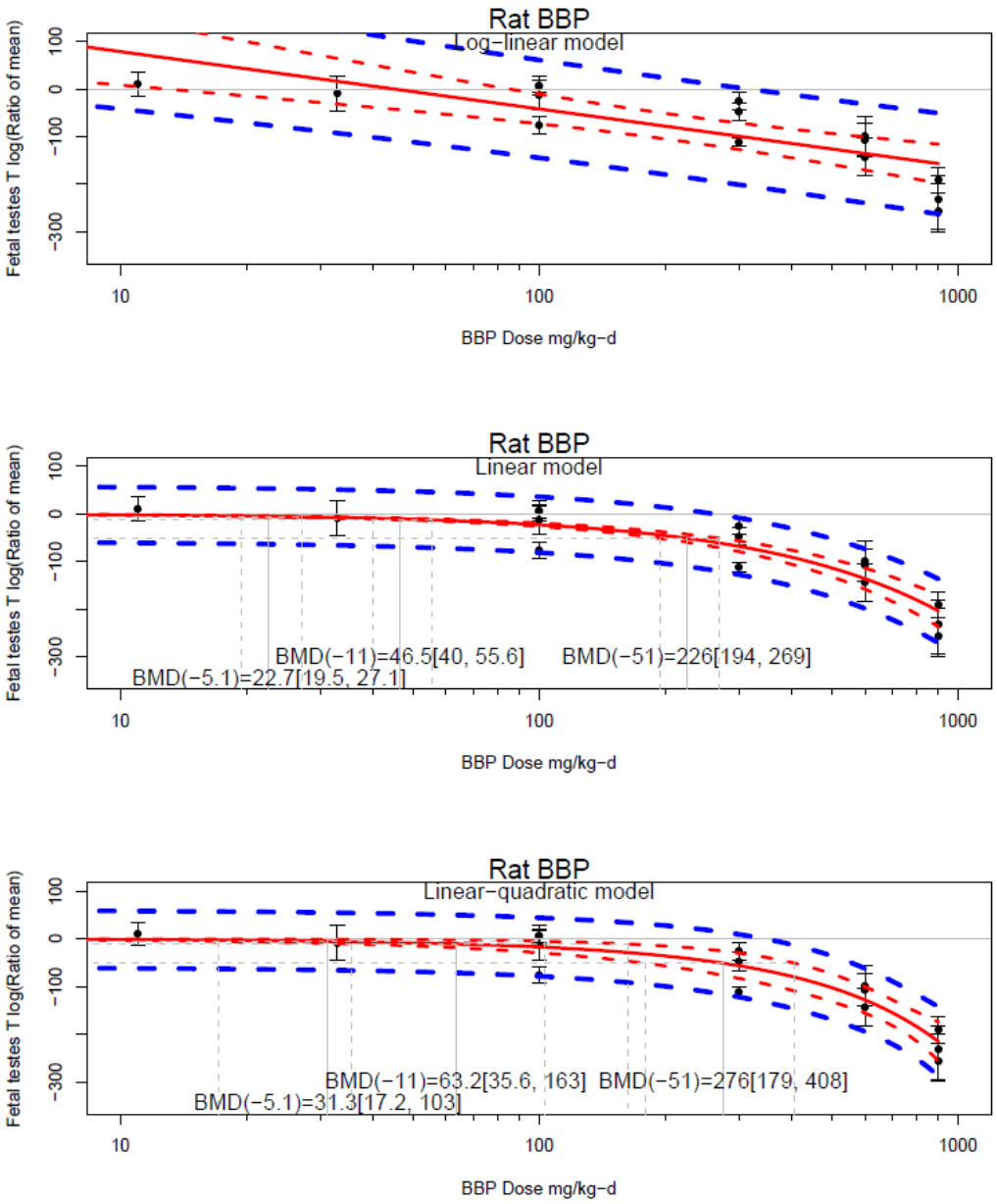
Figure_Apx A-16. Updated Benchmark Dose Estimates from Rat Studies of DIBP and Fetal Testosterone (Metafor Version 4.6.0)

A.5 Butyl Benzyl Phthalate (BBP) – Updated Analysis

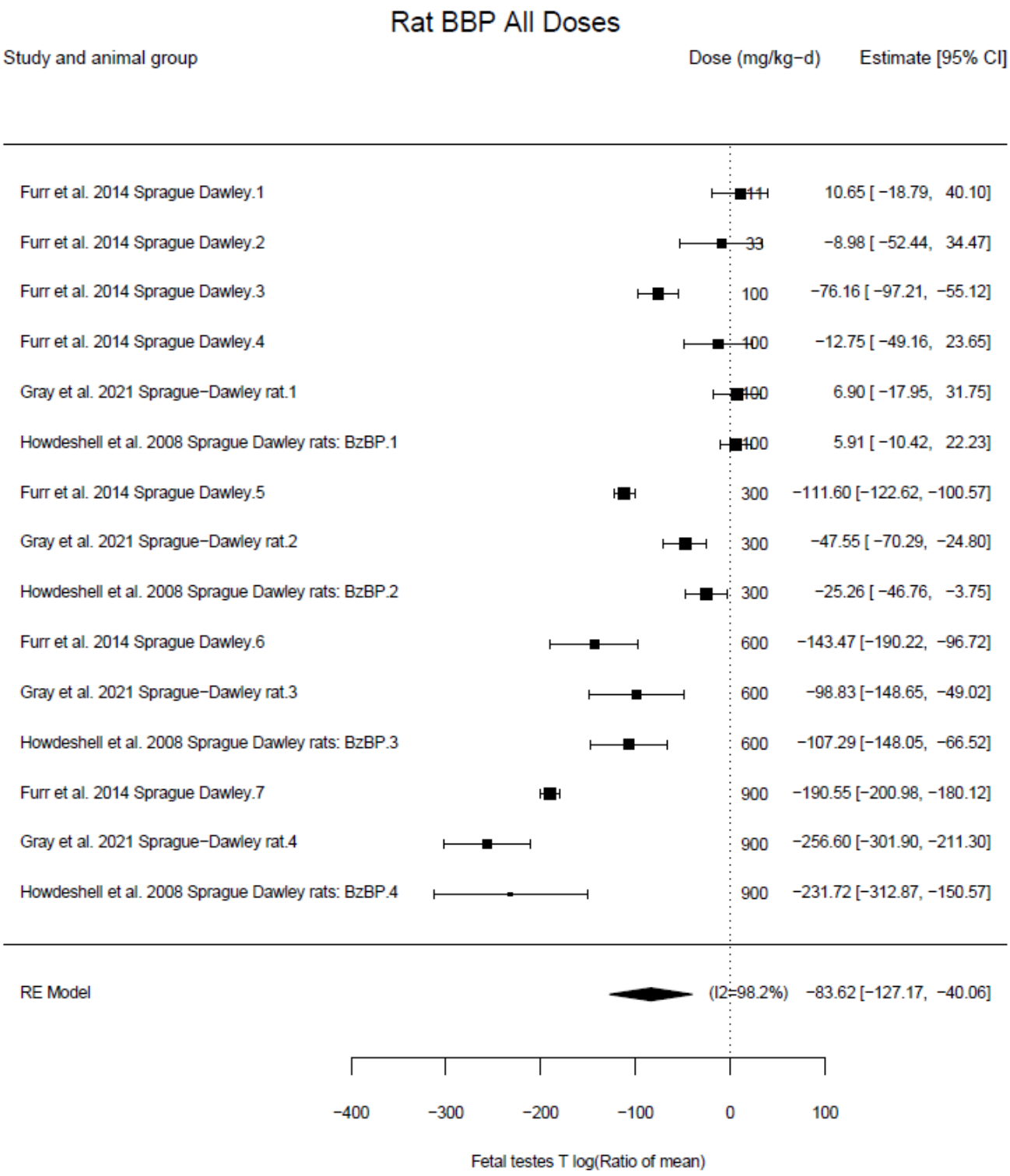


Figure_Apx A-17. Updated Meta-analysis of Studies of BBP and Fetal Testosterone in Rats (Metafor Version 2.0.0)

'Estimate [95% CI]' indicates the estimated effect of BBP on free testes testosterone expressed as the log transformed ratio of means.

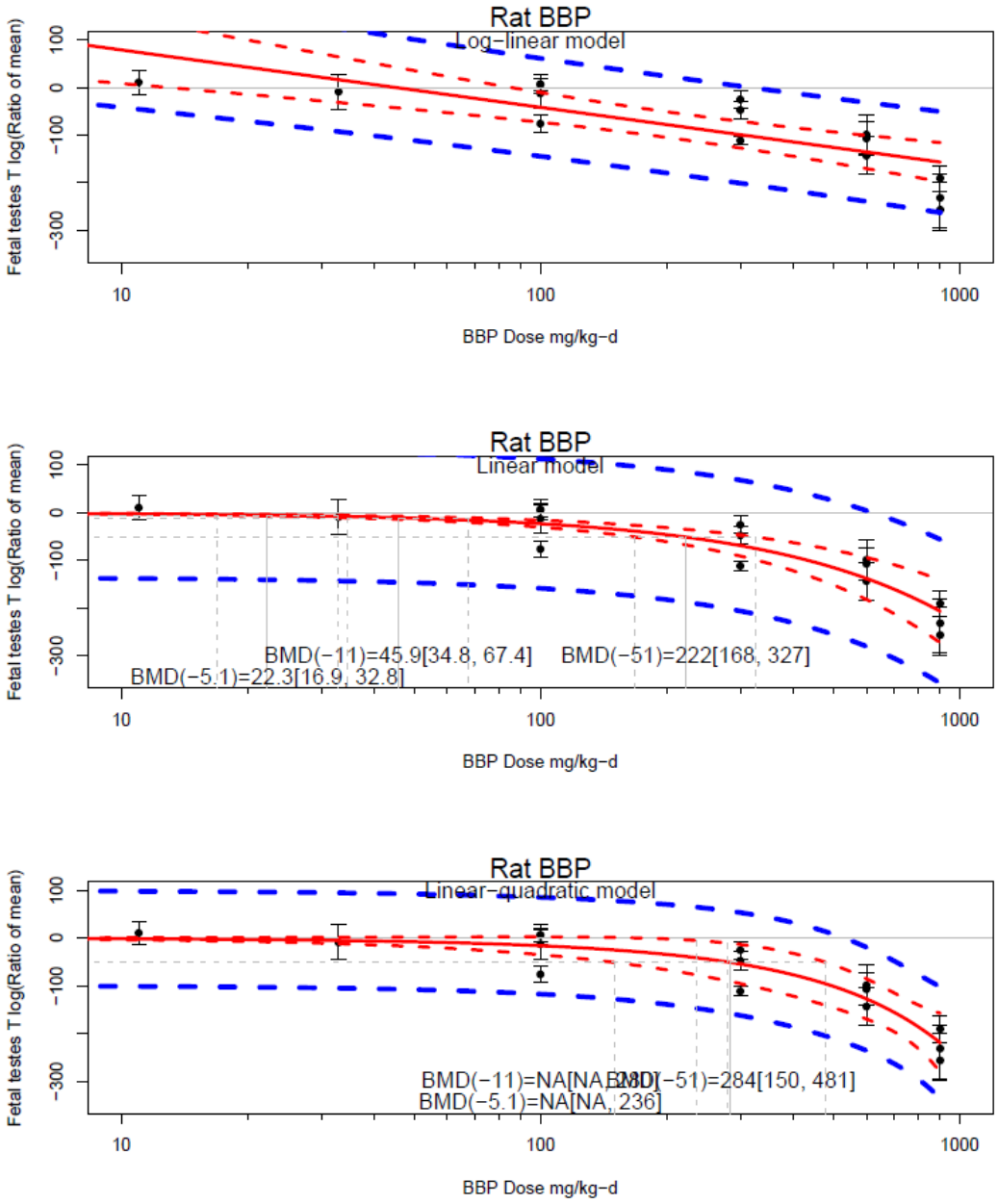


Figure_Apx A-18. Updated Benchmark Dose Estimates from Rat Studies of BBP and Fetal Testosterone (Metafor Version 2.0.0)



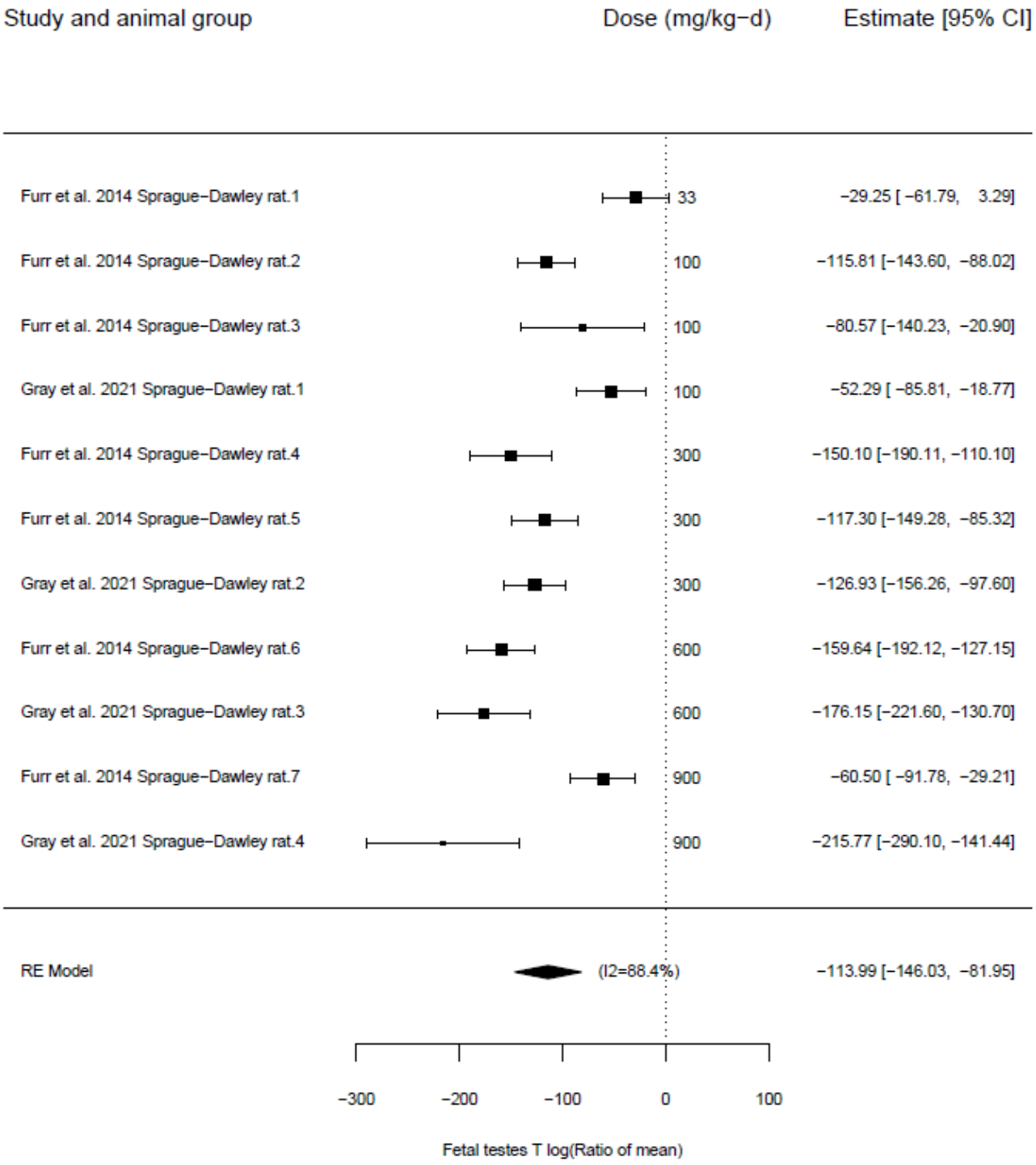
Figure_Apx A-19. Updated Meta-analysis of Studies of BBP and Fetal Testosterone in Rats (Metafor Version 4.6.0)

'Estimate [95% CI]' indicates the estimated effect of BBP on free testes testosterone expressed as the log transformed ratio of means.



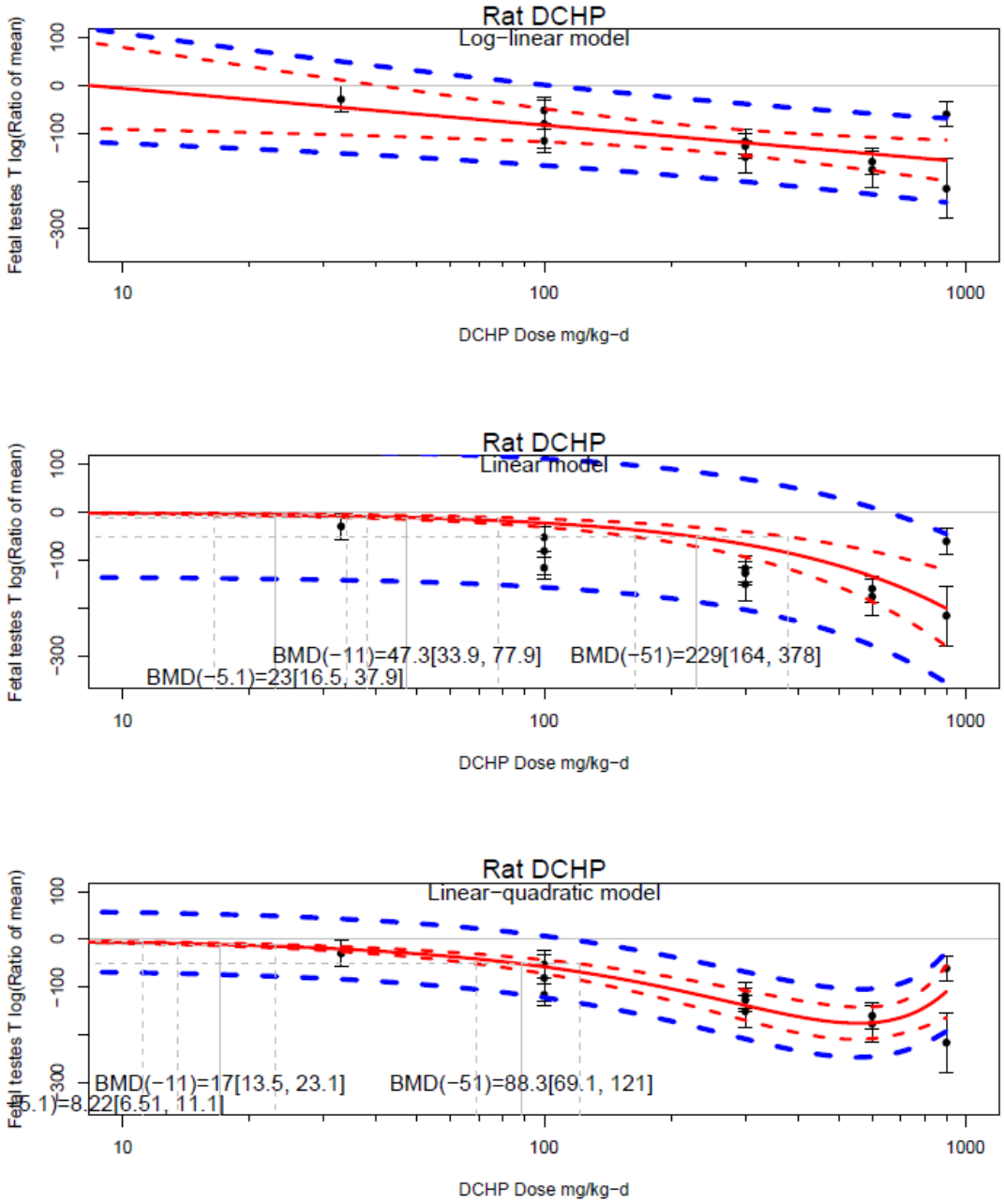
Figure_Apx A-20. Updated Benchmark Dose Estimates from Rat Studies of BBP and Fetal Testosterone (Metafor Version 4.6.0)

A.6 Dicyclohexyl Phthalate (DCHP) – Analysis

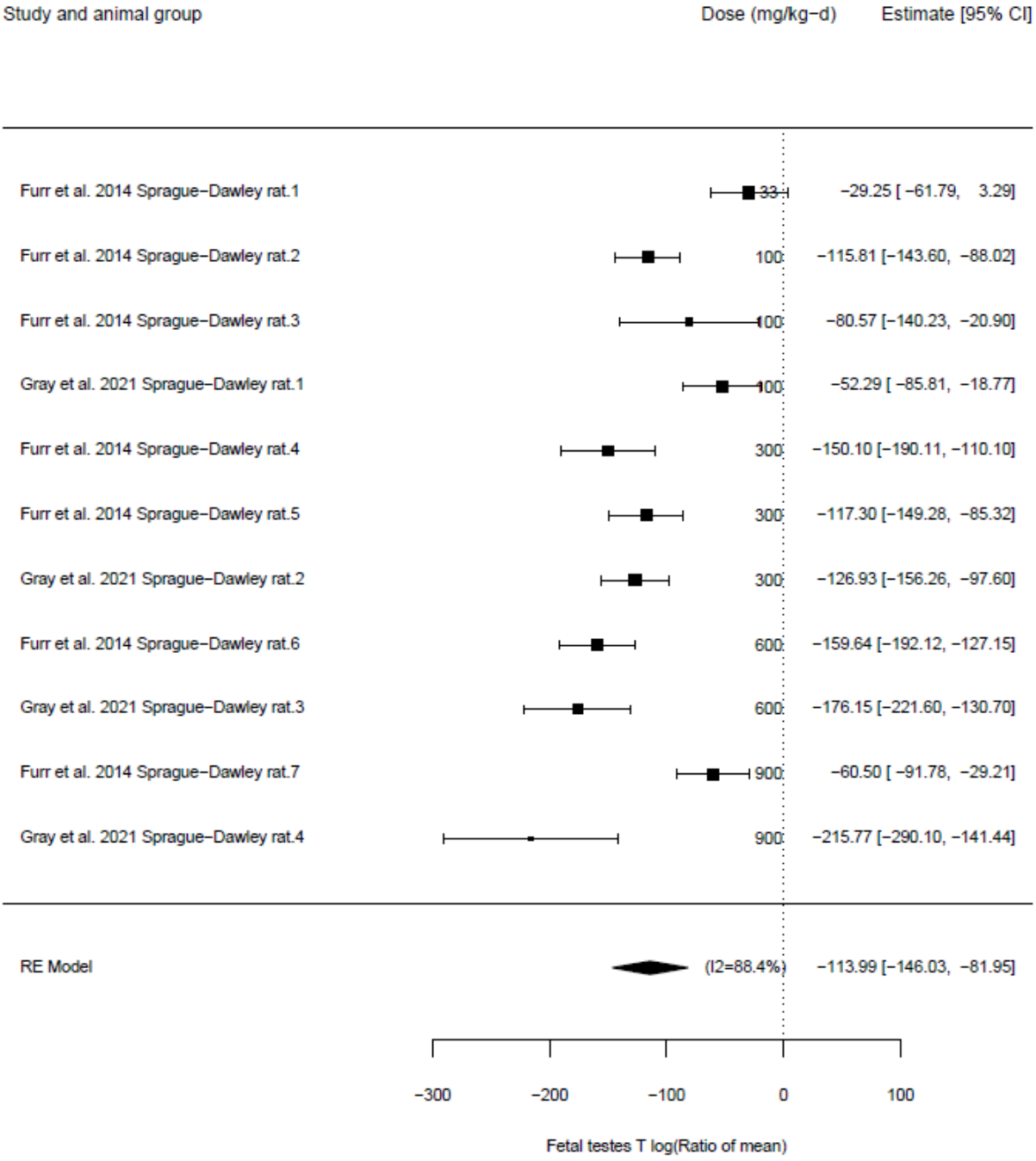


Figure_Apx A-21. Meta-analysis of Studies of DCHP and Fetal Testosterone in Rats (Metafor Version 2.0.0)

'Estimate [95% CI]' indicates the estimated effect of DCHP on free testes testosterone expressed as the log transformed ratio of means.

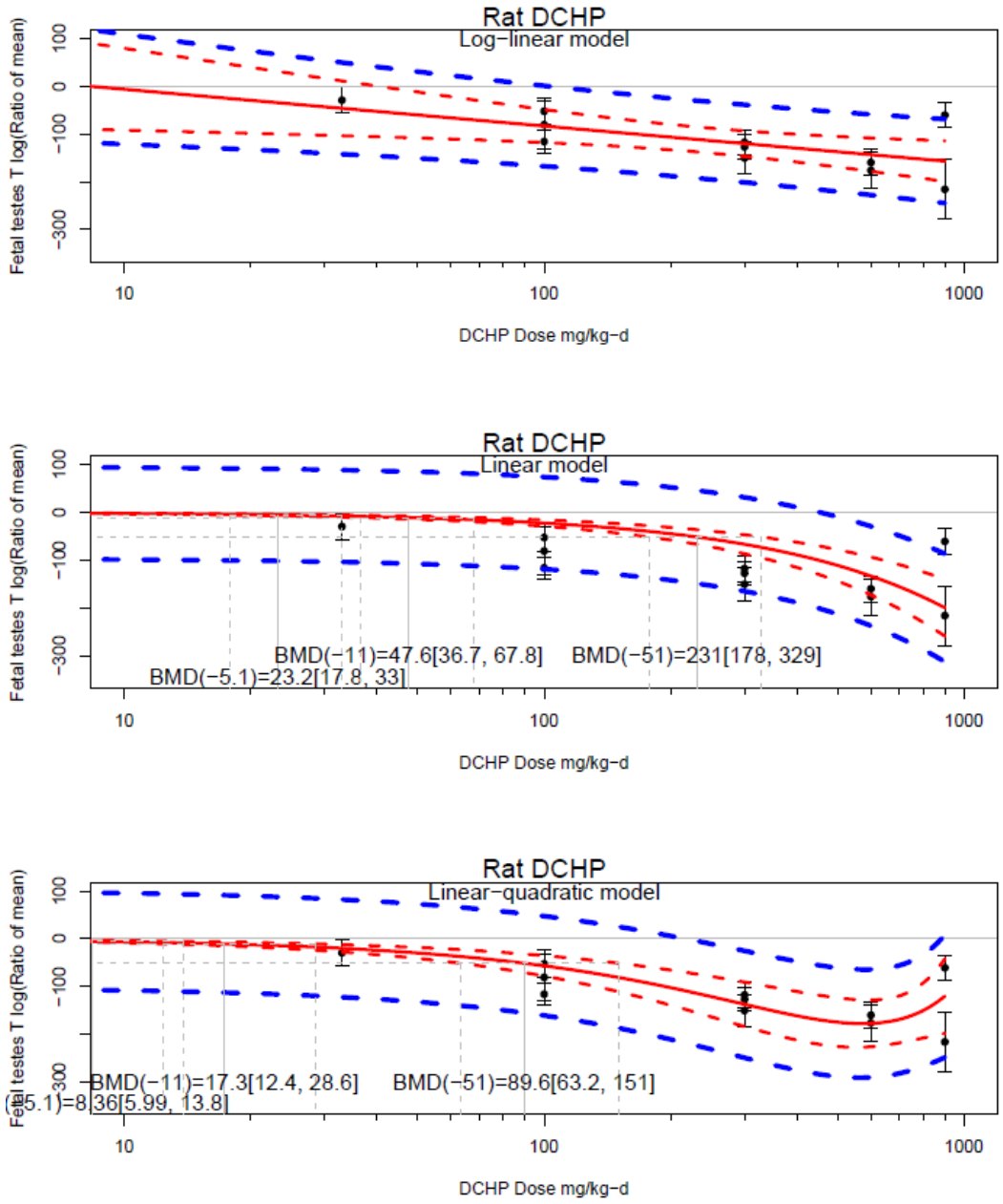


Figure_Apx A-22. Benchmark Dose Estimates from Rat Studies of DCHP and Fetal Testosterone (Metafor Version 2.0.0)



Figure_Apx A-23. Meta-analysis of Studies of DCHP and Fetal Testosterone in Rats (Metafor Version 4.6.0)

'Estimate [95% CI]' indicates the estimated effect of DCHP on free testes testosterone expressed as the log transformed ratio of means.



Figure_Apx A-24. Updated Benchmark Dose Estimates from Rat Studies of DCHP and Fetal Testosterone (Metafor Version 4.6.0)

Appendix B TESTOSTERONE STUDIES CONSIDERED FOR INCLUSION IN META-ANALYSIS

Table_Apx B-1. Summary of Testosterone Studies Considered for Inclusion in DBP Meta-Analysis

Reference	Included in Meta-Analysis by NASEM (2017)?	Included in Updated Meta-Analysis by U.S. EPA (2025)?	Excluded From Meta-Analysis?	Reason For Exclusion From Meta-Analysis
(Furr et al., 2014)	Yes	Yes	No	N/A
(Howdeshell et al., 2008)	Yes	Yes	No	N/A
(Martino-Andrade et al., 2008)	Yes	Yes	No	N/A
(Kuhl et al., 2007)	Yes	Yes	No	N/A
(Struve et al., 2009)	Yes	Yes	No	N/A
(Johnson et al., 2011)	Yes	Yes	No	N/A
(Johnson et al., 2007)	Yes	Yes	No	N/A
(Gray et al., 2021)	No (new study)	Yes	No	N/A
(Clewell et al., 2009)	No	No	Yes	Excluded by NASEM (2017) • N reported as range, not exact value (testosterone reported as mean from 3–4 litters per dose) • Data reported graphically only
(Lehmann et al., 2004)	No	No	Yes	Excluded by NASEM (2017) • N reported as range, not exact value (testosterone reported as average ± SEM from 3–4 rat fetuses from 1–4 dams per dose) • Data reported graphically only
(Mahood et al., 2007)	No	No	Yes	Excluded by NASEM (2017) • N reported as range, not exact value (testosterone reported as mean from 4–6 litters per dose)

Reference	Included in Meta-Analysis by NASEM (2017)?	Included in Updated Meta-Analysis by U.S. EPA (2025)?	Excluded From Meta-Analysis?	Reason For Exclusion From Meta-Analysis
				• Data reported graphically only
(van den Driesche et al., 2012)	No	No	Yes	Excluded by NASEM (2017) • N reported as range, not exact value (testosterone reported as mean from 3–7 litters per dose) • Data reported graphically only
(Li et al., 2015)	No	No	Yes	Excluded by NASEM (2017) • Variance type (standard error or standard deviation) not specified • Data reported graphically only
(Giribabu et al., 2014)	No	No	Yes	Excluded by U.S. EPA • Evaluated serum (not testis) testosterone • Testosterone measured during postnatal (not fetal) lifestage
(Scarano et al., 2010)	No	No	Yes	Excluded by U.S. EPA • Evaluated serum (not testis) testosterone • Testosterone measured during postnatal (not fetal) lifestage
(Kim et al., 2010)	No	No	Yes	Excluded by U.S. EPA • Evaluated serum (not testis) testosterone • Testosterone measured during postnatal (not fetal) lifestage
(Ahmad et al., 2014)	No	No	Yes	Excluded by U.S. EPA • Evaluated serum (not testis) testosterone • Testosterone measured during postnatal (not fetal) lifestage
(Xiao-Feng et al., 2009)	No	No	Yes	Excluded by U.S. EPA • Exposure outside of critical window of development • Evaluated serum (not testis) testosterone

Reference	Included in Meta-Analysis by NASEM (2017)?	Included in Updated Meta-Analysis by U.S. EPA (2025)?	Excluded From Meta-Analysis?	Reason For Exclusion From Meta-Analysis
				• Testosterone measured during postnatal (not fetal) lifestage
(Drake et al., 2009)	No	No	Yes	Excluded by U.S. EPA • N reported as range, not exact value (testosterone reported as litter mean from 1–5 animals per 4–5 litters per group) • Data reported graphically only
(MacLeod et al., 2010)	No	No	Yes	Excluded by U.S. EPA • N reported as range, not exact value (mean reported as being derived from 15–44 intratesticular testosterone values from individual fetuses) • Data reported graphically only
(Mylchreest et al., 2002)	No	No	Yes	Excluded by U.S. EPA • N reported as range, not exact value (mean testosterone values based on N of 4 litters [23–49 fetuses] for control and 5–6 litters [23–49 fetuses] for DBP treatment groups) • Data reported graphically only
(Wilson et al., 2004)	No	No	Yes	Excluded by U.S. EPA • Data reported graphically only
(Howdeshell et al., 2007)	No	No	Yes	Excluded by U.S. EPA • Data reported graphically only
(Spade et al., 2018)	No	No	Yes	Excluded by U.S. EPA • Data reported graphically only
(Gaido et al., 2007)	No	No	Yes	Excluded by U.S. EPA • Evaluated fetal testis testosterone in mice, not rats
(Moody et al., 2013)	No	No	Yes	Excluded by U.S. EPA • Evaluated testosterone in mice, not rats

Reference	Included in Meta-Analysis by NASEM (2017)?	Included in Updated Meta-Analysis by U.S. EPA (2025)?	Excluded From Meta-Analysis?	Reason For Exclusion From Meta-Analysis
				• Exposure outside of critical window of development • Evaluated serum (not testis) testosterone • Testosterone measured during postnatal (not fetal) lifestage
(Li et al., 2023)	No	No	Yes	Excluded by U.S. EPA • Evaluated fetal testosterone in mice, not rats • Evaluated serum (not testis) testosterone
(Higuchi et al., 2003)	No	No	Yes	Excluded by U.S. EPA • Evaluated testosterone in rabbits, not rats • Evaluated serum (not testis) testosterone • Testosterone measured during postnatal (not fetal) lifestage
(McKinnell et al., 2009)	No	No	Yes	Excluded by U.S. EPA • Evaluated testosterone in monkeys, not rats • Evaluated plasma (not testis) testosterone • Testosterone measured during postnatal (not fetal) lifestage

Table_Apx B-2. Summary of Testosterone Studies Considered for Inclusion in DEHP Meta-Analysis

Reference	Included in Meta-Analysis by NASEM (2017)?	Included in Updated Meta-Analysis by U.S. EPA (2025)?	Excluded From Meta-Analysis?	Reason For Exclusion From Meta-Analysis
(Lin et al., 2008)	Yes	Yes	No	N/A
(Furr et al., 2014)	Yes	Yes	No	N/A
(Hannas et al., 2011)	Yes	Yes	No	N/A
(Howdeshell et al., 2008)	Yes	Yes	No	N/A
(Culty et al., 2008)	Yes	Yes	No	N/A
(Martino-Andrade et al., 2008)	Yes	Yes	No	N/A
(Saillenfait et al., 2013)	Yes	Yes	No	N/A
(Gray et al., 2021)	No (new study)	Yes	No	N/A
(Borch et al., 2004)	No	No	Yes	Excluded by NASEM (2017) • N reported as range, not exact value (mean reported as being derived from 6–10, 6–8, or 7-8 litters per dose group, depending upon experiment) • Data reported graphically only
(Borch et al., 2006b)	No	No	Yes	Excluded by NASEM (2017) • N reported as range, not exact value (mean reported as being derived from 5–7 litters per dose group) • Data reported graphically only
(Do et al., 2012)	No	No	Yes	Excluded by NASEM (2017) • Evaluated testosterone in mice, not rats • Evaluated serum (not testis) testosterone
(Klinefelter et al., 2012)	No	No	Yes	Excluded by NASEM (2017) • Fetal testosterone measured after stimulation of testes with luteinizing hormone • Data reported graphically only

Reference	Included in Meta-Analysis by NASEM (2017)?	Included in Updated Meta-Analysis by U.S. EPA (2025)?	Excluded From Meta-Analysis?	Reason For Exclusion From Meta-Analysis
(Vo et al., 2009a)	No	No	Yes	Excluded by NASEM (2017) • Missing group size (N) numbers • Evaluated serum (not testis) testosterone
(Spade et al., 2018)	No	No	Yes	Excluded by U.S. EPA • Data reported graphically only
(Wilson et al., 2004)	No	No	Yes	Excluded by U.S. EPA • Data reported graphically only
(Li et al., 2012)	No	No	Yes	Excluded by U.S. EPA • Exposure outside of critical window of development • Evaluated serum (not testis) testosterone • Testosterone measured during postnatal (not fetal) lifestage • Data reported graphically only
(Vo et al., 2009b)	No	No	Yes	Excluded by U.S. EPA • Exposure outside of critical window of development • Evaluated serum (not testis) testosterone • Testosterone measured during postnatal (not fetal) lifestage • Data reported graphically only
(Gray et al., 2009)	No	No	Yes	Excluded by U.S. EPA • Evaluated serum (not testis) testosterone • Testosterone measured during postnatal (not fetal) lifestage
(Akingbemi et al., 2001)	No	No	Yes	Excluded by U.S. EPA • Evaluated serum (not testis) testosterone • Testosterone measured during postnatal (not fetal) lifestage
(Akingbemi et al., 2004)	No	No	Yes	Excluded by U.S. EPA • Exposure outside of critical window of development • Evaluated serum (not testis) or hormone-stimulated testosterone production • Testosterone measured during postnatal (not fetal) lifestage

Reference	Included in Meta-Analysis by NASEM (2017)?	Included in Updated Meta-Analysis by U.S. EPA (2025)?	Excluded From Meta-Analysis?	Reason For Exclusion From Meta-Analysis
(Lin et al., 2009)	No	No	Yes	Excluded by U.S. EPA • Evaluated serum (not testis) testosterone • Testosterone measured during postnatal (not fetal) lifestage • Data reported graphically only
(Andrade et al., 2006)	No	No	Yes	Excluded by U.S. EPA • Testosterone measured during postnatal (not fetal) lifestage
(Rajagopal et al., 2019)	No	No	Yes	Excluded by U.S. EPA • Evaluated serum (not testis) testosterone • Data reported graphically only • Testosterone measured during postnatal (not fetal) lifestage
(Ge et al., 2007),	No	No	Yes	Excluded by U.S. EPA • Exposure outside of critical window of development • Testosterone measured during postnatal (not fetal) lifestage • Evaluated serum (not testis) testosterone
(Guo et al., 2013)	No	No	Yes	Excluded by U.S. EPA • Exposure outside of critical window of development • Testosterone measured during postnatal (not fetal) lifestage • Evaluated serum (not testis) testosterone
(Barakat et al., 2018)	No	No	Yes	Excluded by U.S. EPA • Evaluated fetal testosterone in mice, not rats • Evaluated serum (not testis) testosterone
(Gaido et al., 2007)	No	No	Yes	Excluded by U.S. EPA • Evaluated fetal testis testosterone in mice, not rats
(Kurahashi et al., 2005)	No	No	Yes	Excluded by U.S. EPA • Exposed via inhalation (not oral) route • Exposure outside of critical window of development • Evaluated serum (not testis) testosterone • Testosterone measured during postnatal (not fetal) lifestage

Reference	Included in Meta-Analysis by NASEM (2017)?	Included in Updated Meta-Analysis by U.S. EPA (2025)?	Excluded From Meta-Analysis?	Reason For Exclusion From Meta-Analysis
(Ma et al., 2006)	No	No	Yes	Excluded by U.S. EPA • Exposed via inhalation (not oral) route • Exposure outside of critical window of development • Evaluated serum (not testis) testosterone • Testosterone measured during postnatal (not fetal) lifestage

Table_Apx B-3. Summary of Testosterone Studies Considered for Inclusion in DIBP Meta-Analysis

Reference	Included in Meta-Analysis by NASEM (2017)?	Included in Updated Meta-Analysis by U.S. EPA (2025)?	Excluded From Meta-Analysis?	Reason For Exclusion From Meta-Analysis
(Hannas et al., 2011)	Yes	Yes	No	N/A
(Howdeshell et al., 2008)	Yes	Yes	No	N/A
(Gray et al., 2021)	No (new study)	Yes	No	N/A
(Saillenfait et al., 2017)	No	No	Yes	Excluded by U.S. EPA • Data reported graphically only
(Borch et al., 2006a)	No	No	Yes	Excluded by U.S. EPA • N reported as range, not exact value (mean reported as being derived from N of 5–6 per dose group) • Data reported graphically only
(Pan et al., 2017)	No	No	Yes	Excluded by U.S. EPA • Evaluated testosterone in mice, not rats • Testosterone measured during postnatal (not fetal) lifestage • Data reported graphically only
(Wang et al., 2017)	No	No	Yes	Excluded by U.S. EPA • Evaluated testosterone in mice, not rats • Exposure outside of critical window of development • Testosterone measured during postnatal (not fetal) lifestage • Data reported graphically only

Table_Apx B-4. Summary of Testosterone Studies Considered for Inclusion in BBP Meta-Analysis

Reference	Included in Meta-Analysis by NASEM (2017)?	Included in Updated Meta-Analysis by U.S. EPA (2025)?	Excluded From Meta-Analysis?	Reason For Exclusion From Meta-Analysis
(Howdeshell et al., 2008)	Yes	Yes	No	N/A
(Furr et al., 2014)	Yes	Yes	No	N/A
(Gray et al., 2021)	No (new study)	Yes	No	N/A
(Nagao et al., 2000)	No	No	Yes	Excluded by U.S. EPA • Evaluated serum (not testis) testosterone • Testosterone measured during postnatal (not fetal) lifestage
(Aso et al., 2005)	No	No	Yes	Excluded by U.S. EPA • Evaluated serum (not testis) testosterone • Testosterone measured during postnatal (not fetal) lifestage
(Ahmad et al., 2014)	No	No	Yes	Excluded by U.S. EPA • Evaluated serum (not testis) testosterone • Testosterone measured during postnatal (not fetal) lifestage
(Spade et al., 2018)	No	No	Yes	Excluded by U.S. EPA • Data reported graphically only
(Wilson et al., 2004)	No	No	Yes	Excluded by U.S. EPA • Data reported graphically only
(Schmitt et al., 2016)	No	No	Yes	Excluded by U.S. EPA • Evaluated testosterone in mice, not rats • Evaluated serum (not testis) testosterone • Testosterone measured during postnatal (not fetal) lifestage

Table_Apx B-5. Summary of Testosterone Studies Considered for Inclusion in DCHP Meta-Analysis

Reference	Included in Meta-Analysis by NASEM (2017)?	Included in Updated Meta-Analysis by U.S. EPA (2025)?	Excluded From Meta-Analysis?	Reason For Exclusion From Meta-Analysis
(Furr et al., 2014)	N/A [a]	Yes	No	N/A
(Gray et al., 2021)	N/A [a]	Yes	No	N/A
(Ahbab and Barlas, 2013)	N/A [a]	No	Yes	Excluded by U.S. EPA • Evaluated serum (not testis) testosterone • Testosterone measured during postnatal (not fetal) lifestage
(Hoshino et al., 2005)	N/A [a]	No	Yes	Excluded by U.S. EPA • Evaluated serum (not testis) testosterone • Testosterone measured during postnatal (not fetal) lifestage
(Li et al., 2016)	N/A [a]	No	Yes	Excluded by U.S. EPA • Testosterone measured during postnatal (not fetal) lifestage
(Ahbab and Barlas, 2015)	N/A [a]	No	Yes	Excluded by U.S. EPA • Evaluated serum (not testis) testosterone
(Lv et al., 2019)	N/A [a]	No	Yes	Excluded by U.S. EPA • Evaluated serum (not testis) testosterone

[a] Not applicable. DCHP was not included in the NASEM (2017) meta-analysis.